Genomic Analysis of Cancer Heterogeneity and Oncogenic Mechanisms

by

Xiaolei Jiang

University Program in Genetics and Genomics
Duke University

Date:_______________________
Approved:

____
Jen-Tsan Ashley Chi, Supervisor

____
Sandeep Dave

____
Shiaowen David Hsu

____
David Kirsch

____
Allison Ashley-Koch

A dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in the
University Program in Genetics and Genomics in the Graduate School
of Duke University

2014

ABSTRACT

Genomic Analysis of Cancer Heterogeneity and Oncogenic Mechanisms

by

Xiaolei Jiang

University Program in Genetics and Genomics
Duke University

Date:_______________________
Approved:

____
Jen-Tsan Ashley Chi, Supervisor

____
Sandeep Dave

____
Shiaowen David Hsu

____
David Kirsch

____
Allison Ashley-Koch

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the University Program in
Genetics and Genomics in the Graduate School
of Duke University

2014

Abstract

The development of cancer is a process by which an accumulation of genetic changes leads to uncontrolled replication of cells. Since the process of mutation is random, the set of alterations that occur and accumulate during tumorigenesis in one individual is different from that of another. These genetic differences drive tumor heterogeneity. One of the first technologies used to explore genome-wide heterogeneity was the microarray, which can be used to measure the expression of tens of thousands of genes. By exploring differences in expression of not just single genes, but groups of genes that may be altered in one set of tumors compared to another, researchers were able to classify subtypes of cancer that had relevance in disease aggressiveness, treatment, and prognosis. Furthermore, by looking at genome-wide patterns of expression, it is possible to identify specific oncogenic pathways that are activated and critical in driving tumor cell survival, growth, or metastasis. My research utilizes the patterns of expression derived from microarray analyses to study tumor heterogeneity, particularly in response to targeted cancer therapy, and mechanisms of cell death following oncogenic deregulation.

One of the cancer types that has been explored through expression array analysis is B-cell lymphoma. Human aggressive B-cell non-Hodgkin lymphomas (NHL) encompass the continuum between Burkitt lymphoma (BL) and diffuse large B-cell

lymphoma (DLBCL), and display considerable clinical and biologic heterogeneity, most notably related to therapy response. We previously showed that lymphomas arising in the *Eμ-Myc* transgenic mouse are heterogeneous, mirroring genomic differences between BL and DLBCL. Given the clinical heterogeneity in NHL and the need to develop strategies to match therapeutics with discrete forms of disease, we investigated the extent to which genomic variation in the *Eμ-Myc* model predicts response to therapy. We used genomic analyses to classify *Eμ-Myc* lymphomas, link *Eμ-Myc* lymphomas with NHL subtypes, and identify lymphomas with predicted resistance to conventional and NF-κB targeted therapies. Experimental evaluation of these predictions links genomic profiles with distinct outcomes to conventional and targeted therapies in the *Eμ-Myc* model, and establishes a framework to test novel targeted therapies or combination therapies in specific genomically-defined lymphoma subgroups. In turn, this will rationally inform the design of new treatment options for aggressive human NHL.

The second aspect of my thesis looks at the mechanisms of apoptosis following oncogene deregulation. The Rb-E2F pathway is a critical oncogenic pathway that is frequently mutated in cancers. Alterations in the pathway affect genome-wide expression in the cell, which in turn lead to deregulation of the cell cycle. The E2F1 transcription factor regulates cell proliferation and apoptosis through the control of a considerable variety of target genes. Previous work has detailed the role of other

transcription factors that cooperate with E2F to mediate the specificity of E2F function. In this work, we identify the NF-YB transcription factor as a novel direct E2F1 target. Genome-wide expression analysis of the effects of NFYB knockdown on E2F1-mediated transcription identified a large group of genes that are co-regulated by E2F1 and NFYB. We also provide evidence that knockdown of NFYB enhances E2F1-induced apoptosis, suggesting a pro-survival function of the NFYB/E2F1 joint transcriptional program. Bioinformatic analysis suggests that deregulation of these NFY-dependent E2F1 target genes might play a role in sarcomagenesis as well as drug resistance.

Taken together, these studies highlight the importance and power of analyzing genome-wide patterns of expression in investigating cancer heterogeneity, its ability to help predict treatment response, and its role in discovering the mechanisms behind the consequences of gene deregulation.

Dedication

This work is dedicated to my amazing wife, Su, and our wonderful son, Aiden, as well as my parents, Jinchun Jiang and Aiqing Wang.

Contents

List of Tables

List of Figures

Acknowledgements

The work described here has been made possible by tremendous contributions from scientific minds far greater than my own. It has been a tremendous learning experience to see their brilliant minds at work. This starts with my two mentors, Joe Nevins and Ashley Chi. From the beginning of my graduate career, Joe helped me to find my path, guide me through my projects, and focus whenever I lost sight of the purpose of my experiments. Being able to see the way he thinks about problems and understands complex issues is something I have always been amazed at and learned from. His dedication to the people in his lab, even while going through his own struggles, is something I hope to emulate. During my first few years of graduate school, when Ashley was one of my committee members, he often would check on my progress and give me feedback and advice. It was very fitting that I joined his lab to finish my projects after Joe's retirement, as he has helped to guide me through perhaps the most difficult times in my later graduate school years. I would also like to thank my committee members, both past and present, Sandeep Dave, David Hsu, David Kirsch, Allison Ashley-Koch, and Joe Lucas, for their scientific feedback and advice as well as encouragement.

Along with my advisors, the individuals in their labs have also been hugely important as teachers, colleagues, and friends. I was fortunate to work with two wonderful people, Rachel and Daphne, as part of “Team Lymphoma”. Scientifically, I learned so much from them. But even more than that, I found two friends who I enjoyed spending hours with at the mouse room, at the flow cytometer, in the tissue culture room, and at the bench. I have also had a chance to work closely with Igor, who I have a great deal of respect for as a scientist and as a person. Whenever I had trouble designing an experiment, it always seemed like Igor knew exactly what needed to be done. Through our discussions, I have learned how to think more effectively as a scientist and focus on finding the important results amongst a sea of data. Laszlo has also been a tremendous colleague and friend since the beginning. His jokes helped lift my spirits and he was always there to provide encouragement and recognition for my efforts. Kaye was someone who I was actually scared to talk to at first given her stern demeanor. Yet has I got to know her, I discovered she was an incredibly kind individual always eager to help me, whether it be organizing committee meetings or making sure an order went through. Dawn was incredibly helpful in setting up some of my projects and always providing the lab with the materials to make science happen. Jason and Heather, as fellow graduate students, were a source of inspiration as well as great friends to talk to on both a professional and personal level.

Changing labs can certainly be a daunting experience, but the people in Ashley’s lab made the transition incredibly easy. They are a great group of dedicated individuals who I am fortunate to have known. Chien-Kuang, through our conversations about life more than science, has become a great friend whose positivity cheers me up. Her dedication to her labmates is evident through her remembering everyone's birthday and always having a birthday cake ready. Melissa, a.k.a. "Mel", has helped me resolve many experimental problems and is also someone who I can share my frustration with about Westerns. Jianli has become a great friend and incredibly helpful in helping me order everything I needed for my experiments. She also gives me a chance to practice my Chinese at the workplace.

While my labmates have made my graduate career a success, I would not have even been in graduate school without the support of my family. My dad came to the US from a poor rural village in China to create a better life for his family. My mom came to the US with no knowledge of English. Together, they worked incredibly hard to make sure I had every opportunity to succeed. They always encouraged me to pursue my education first. My PhD is as much theirs as it is mine. At the beginning of graduate school, I was fortunate to meet the person I will spend the rest of my life with. I cannot imagine what the past six years would have been like without my wife, Su. No matter how difficult a day I had at the lab, knowing that I could go home and see her kept me going through the greatest frustrations. Watching her work ethic helped push me to work harder. If I had to miss dinner because of harvesting cells at specific time points, she would cook and bring food to me to make sure I would not go hungry. We have now grown into a small family, as we welcomed our son, Aiden, earlier this year. It has been an absolutely joy to watch him grow and learn. I am thankful every day that I get to see the two of them and I am incredibly excited about what the future holds for us.

1. Introduction

1.1 Molecular heterogeneity of cancer

In their two seminal papers, Douglas Hanahan and Robert Weinberg proposed eight hallmark capabilities that all cancers needed: 1. Sustained proliferative signaling, 2. Resistance to cell death, 3. Evasion of growth suppressors, 4. Replicative immortality, 5. Induction of angiogenesis, 6. Activating invasion and metastasis, 7. Genome instability and mutation, and 8. Tumor-promoting inflammation, as well as two emerging hallmarks, deregulation of cellular energetics and avoidance of immune destruction (1, 2). The development of cancer is the process of acquiring stochastic mutations, some of which will cause deregulation of various oncogenic pathways that confer tumorigenic capabilities that include uncontrolled proliferation and resistance to apoptosis signals, as mentioned in within the cancer hallmarks. Hence, the mutations that define the oncogenic process in one tumor will most likely be very different from the mutations that confer oncogenesis in another. This suggests that cancer is heterogeneous with differing sets of mutations leading to tumor characteristics that are unique within a group of tumors with a similar mutational profile.

Ideally, each tumor could be sequenced and its set of defining mutations determined. However, the cost of genome wide sequencing has been prohibitively expensive until recent years (3). Although it is now financially feasible to sequence the genetic material of tumors, a different technology, the DNA microarray, has already been extensively used for many years to study the differences in gene expression between different tumors and normal cells, leading to greater understanding of gene deregulation and tumor heterogeneity.

1.1.1 DNA microarrays

DNA microarray technology is based on both the dot blot, in which fragmented DNA from a sample of interest is transferred to a membrane, and then probed with a known DNA sequence of interest that is radiolabeled or fluorophore-labeled (4). Microarrays apply this technique to hundreds or thousands of DNA probes so that the presence and or quantity of many sequences of interest can be determined at the same time. Their first use was documented in 1982, when cDNA synthesized from induced mouse colon carcinoma RNA were cloned into *E. coli*. Clones were grown on nitrocellulose filters and screened with radiolabelled cDNA synthesized from polyadenylic acid-containing cytoplasmic RNA (messenger RNA) extracted from normal mouse colon, liver, kidney, or colon tumors (5). The relative brightness of the resultant spots on the filters was assessed to determine the relative expression of the clone sequences in the tumor and normal tissue samples. Later studies by the same group expanded the number of cloned sequences and showed differences between normal colon tissue, benign colonic adenomas, and colon carcinomas (6). This was also the first instance in which computer-based scanning and image-processing were used, which, given the increasing number of clones, became necessary for efficient data collection. A few years later, the same group used arrays with even more sequences and found a set of thirty clones, possibly the first multi-gene signature, that differentiated normal colonic mucosa of low risk individuals from normal colonic mucosa of high risk individuals who had familial adenomatous polyposis or hereditary nonpolyposis colon cancer (7).

Inter-tumor cancer heterogeneity has been recognized by the scientific community for a number of decades and certainly before the use of microarrays. However, microarrays allowed for the first time the determination of tumor gene expression on a genome-wide level. Also, with the miniaturization and development of spotted arrays on glass slides, in which small microscopic drops of oligonucleotide probes could be positioned in very close proximity, it was now possible to assay more genes with less sample material (8). In 1999, Eric Lander's group at Harvard, using commercial arrays, showed that expression arrays can be used to discover distinctions between acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL), two diseases which at that point were difficult to accurately differentiate clinically (9). The following year, Brown and Staudt's group at Stanford, using their own high density arrays, defined novel germinal center B-cell (GCB) and activated B-cell (ABC) subtypes of diffuse large B cell lymphoma, a disease which had previously been treated clinically as a single entity (10). These subtypes reflected the lymphomas' cell of origin and demonstrated drastically different clinical prognoses. That same year, they assessed human breast cancer and were able to classify them into several subtypes based on the differences in their pattern of gene expression as well as show differences in clinical outcomes (11). Subsequent studies have used similar methods to "profile" lung cancer (12, 13), prostate cancer (14), and others. By studying the patterns of gene expression that define specific "molecular" subtypes of various cancers, it was feasible to look at what oncogenic pathways might be affected by the set of mutations within a certain tumor.

1.1.2 Pathway-based heterogeneity

Since gene expression profiling of cancers began, researchers were also looking at how to classify new tumors into the identified subgroups. When Eric Lander's group showed through gene expression that AML and ALL can be distinguished, they also used the genes whose expression patterns differentiated the two cancers to generate a predictive model in which new tumors could be assigned to one group or another (9). Hence, the concept of a gene signature, a set of genes in which expression is correlated with one state or another, was applied to studying cancer. Subsequently, the same research group generated a gene expression signature for metastasis that was derived from a comparison of primary adenocarcinomas and metastases (15). This signature was further demonstrated to stratify tumors which were likely to be associated with metastasis and hence showed poor clinical prognosis. A group at the Netherlands Cancer Institute under René Bernards generated a signature that predicted the likelihood of breast cancer metastasis and was also linked to clinical outcomes (16, 17). Signatures were also applied to study the patterns of gene expression changes following oncogene deregulation. By assessing gene expression patterns across hundreds of tumors, the Harvard group extracted a signature for cyclin D1 overexpression (18).

Leveraging the idea of the expression of entire pathways being altered following oncogenic alterations, researchers sought to mimic the effect of the activation by particular oncogenic mutations or of oncogenic pathways via genetic manipulation or growth factors. The Nevins group at Duke used adenoviruses containing an oncogene (Myc, Ras, E2F1, E2F2, E2F3) or a negative control and compared the patterns of expression. The resulting expression changes should represent those genes whose expression varied across samples in a way that segregated oncogene-deregulated from control samples (19). These gene signatures were able to accurately calculate when a pathway was activated during the cell cycle based on the patterns of gene expression at that point in the cycle. This work was further expanded with additional oncogenic pathways such as Src and β-catenin. More importantly, the signatures were able to accurately predict the likelihood and degree to which an oncogenic pathway was activated in tumors (20). As an example, MYC-driven mouse tumors from MMTV-Myc mice showed high Myc pathway activation compared to HER2 driven, Rb null, or wild type tissue. Similarly, Rb null tumors showed the highest E2F3 pathway activation compared to other tumor types, consistent with Rb inhibition of E2Fs. In non-small cell lung cancer, the Ras signature was significantly higher in tumors bearing a Ras activating mutation. Furthermore, profiling based on patterns of pathway activation rather than raw expression data can not only separate tumors into subgroups, but also provide information on pathways likely to be active and driving growth of the tumor (21).

Similar pathway signature-based approaches have be also used to analyze non-genetic factors, such as tumor microenvironmental stresses, studied by the Chi lab at Duke and other groups. Several examples include gene signatures of cell-type specific hypoxia response that delineate a subset of tumors with strong hypoxia response and poor outcomes in several types of cancers (22-26). The Chi group also developed a signature for lactic acidosis (LA) response, distinct from hypoxia response, that recognizes tumors with more favorable outcomes due to its ability to inhibit Akt and glycolysis and trigger starvation responses (27, 28). One critical mediator of LA response is TXNIP (28), which inhibits glycolysis and exhibits many anti-tumor properties (29-32).

The gene signature approach enables the integrative analysis of non-genetic stress pathways with oncogenic signaling pathways. This has allowed the identify dramatic metabolic differences between basal and luminal types of breast tumors. Basal-type cells have an exaggerated hypoxia response due to high HIF-1α/HIF-2α mRNA levels (33) and are addicted to glutamine and thus susceptible to glutamine deprivation (34). Results from the Chi group also suggest that sub-signatures of the hypoxia/LA responses appear to be driven by DNA copy number alterations (CNAs). These CNAs significantly associate with the hypoxia/LA response in tumors and may regulate and possibly be selected by these stresses (35). For example, a CNA that shows significant negative association with the hypoxia response contains *FH* and *EGLN3*, which code for proteins that negatively regulate HIF-1α (35), a transcription factor essential in cellular response to hypoxia. Similar to oncogenic signatures defined by the deregulation of oncogenes and tumor suppressors, the study of the tumor microenvironment provides another layer with which to examine tumor heterogeneity and the driving forces behind oncogenesis and tumor survival.

Importantly, with the development of cancer therapies that target specific oncogenic proteins rather than killing dividing cells, as is the case with chemotherapy or radiation therapy, it has become important to understand what pathways are important in the survival and growth of a tumor. This has made microarrays and expression signatures an important tool for identifying which cancer patients would be sensitive to which therapies.

1.1.3 Targeted cancer therapeutics

Targeted therapy is an important and relatively new treatment modality that recognizes and treats the heterogeneity inherent in cancer. In contrast to the traditional cytotoxic chemotherapy, which uses toxins to preferentially kill rapidly dividing cells, targeted therapies inhibit specific protein functions within a cancer cell, particularly those that have been deregulated as a result of mutations known to drive oncogenesis or promote tumor survival. They are somewhat similar to hormonal therapy, a modality in which the exogenous administration of hormones, in particular steroid hormones, is used against cancers that rely on a hormone pathway for growth. There are several classes of targeted therapies including tyrosine kinase inhibitors, serine/threonine kinase inhibitors, and monoclonal antibodies.

Tyrosine kinase inhibitors (TKI) are small molecules that act against the phosphorylation activity of kinases that phosphorylate tyrosine residues on target proteins. Deregulated tyrosine kinase activity has been recognized as a key driver of oncogenesis and cancer survival and proliferation. In fact, some cancers, such as chronic myelogenous leukemia, based on their response to tyrosine kinase inhibitors, seem to be strongly if not completely dependent on a specific deregulated tyrosine kinase. This dependence was termed “oncogene addition” (36).

TKIs were first introduced clinically with imatinib (Gleevec) to treat chronic myelogenous leukemia. In the majority of CML tumors, there is a reciprocal translocation between chromosome 9 and 22, called the Philadelphia chromosome, which juxtaposes ABL1, a proto-oncogenic tyrosine kinase, with a portion of breakpoint cluster region protein (BCR) (37). This creates a tyrosine kinase that is constitutively active. Upon treatment with Gleevec, CML patients with the BCR-ABL fusion protein are effectively cured of their disease. Since then, over a dozen other TKIs have been introduced clinically. Gefinitib (Iressa) and erlotinib (Tarceva) target EGFR and are used in the treatment of non-small cell lung cancer. Lapatinib is a dual EGFR/HER2 inhibitor used in the treatment of breast cancer. TKIs that target VEGFR, important in vasculogenesis and angiogenesis (38), SRC, and MET, as well as other tyrosine kinases, are used to treat various cancers and other diseases driven by tyrokine kinase deregulation (reviewed in (39)). Many tyrosine kinases now have inhibitors at the research and development or clinical trial phase of development, such as ALK, which is translocated in many anaplastic large-cell lymphomas as well as some non-small cell lung cancers, and ROS, also found to be rearranged in some non-small cell lung cancers (40).

Serine/Threonine kinase inhibitors target the function of kinases that phosphorylate serine or threonine. Serine/threonine kinases (STK) themselves play important roles in regulating cell proliferation, growth, survival and protein synthesis and have been found to be mutated or deregulated in cancers (41, 42). STK inhibitors are not as numerous clinically as TKIs. Currently, there are small molecule inhibitors targeting BRAF, found to be mutated in some human cancers (42), MEK, and mTOR approved in the treatment of renal cell carcinoma, breast cancer, and melanoma as well as other non-cancer diseases. Similar to TKIs, numerous small molecules are in drug development pipelines targeting many of the kinases within the family.

Monoclonal antibody therapy is unique from kinase inhibitors in that it takes advantage of biological molecules, in this case antibodies that can stimulate a patient's immune system to attack cells bound by those antibodies. The technology originated with the development of hybridomas, hybrid cell lines fusing a B cell that produces antibodies with a myeloma (43). This created an immortal cell line that produced monoclonal antibodies. With the subsequent development chimeric, humanized, and human antibodies that can be produced by transgenic mice or phage display libraries, the resulting antibodies had greater serum half-life and binding affinity to their human targets, as well as reduced risk of allergic reactions (44-46).

The first definitive evidence showing that monoclonal antibodies could work was their use to target cells that had been transformed by Her2 (47-49), a receptor tyrosine kinase responsible for activating various signaling pathways that promote cell proliferation and inhibit apoptosis. HER2 is amplified or overexpressed in approximately 20% of breast cancers (50). Since then, over thirty have been approved for diseases ranging from autoimmune disease to asthma. In 1997, rituximab became the first antibody approved to treat human cancer. Targeting CD20, a B cell antigen expressed beginning at the pro-B phase and increasing in concentration as the cell matures (51), the chimeric antibody was indicated for non-Hodgkin lymphoma. Along with cyclophosphamide, doxorubicin, vincristine, and prednisone, it has become the standard therapy for treatment of most non-Hodgkin lymphomas. This was followed by trastuzumab, a humanized HER2 antibody for breast cancer treatment, the first to be approved for the treatment of solid tumors. Because HER2 is overexpressed in only a subset of cancers, companion diagnostic tests based on immunohistochemistry were established to determine if a breast tumor had an overabundance of HER2 receptors on its cell membranes. Because its efficacy and use in a subset of breast tumors, trastuzumab became the poster child for the promise of targeted therapeutics. Subsequent cancer treating monoclonal therapies have been indicated for use in acute myelogenous leukemia (CD33), chronic lymphocytic leukemia (CD52), colorectal and non-small cell lung cancer (VEGF and EGFR), head and neck cancer (EGFR), and melanoma (CTLA-4).

With the ability of monoclonal antibodies to bind cancer cells that overexpress their targets, having a cytotoxic drug conjugated to the antibody would allow the targeted delivery of chemotherapies directly to cancer cells. Anti-body drug conjugates (ADC) have made this a reality. They not only allow for chemotherapy discrimination between normal and malignant tissues, but also decrease the potentially damaging side effects associated with systemic exposure to cytotoxic agents and increase the dosage that can be used, enhancing chemotherapy effectiveness. More than two dozen ADCs are in various stages of development, with two, a CD30 antibody-antitubulin conjugate (brentuximab vedotin) approved for treatment of hogdkin’s lymphoma and anaplastic large cell lymphoma and Trastuzumab emtansine, a HER2-antitubulin conjugate indicated for treatment of HER2-positive breast cancer (52).

As a standalone therapy or in combination with existing hormonal or cytotoxic therapies, targeted therapies have accounted for many of the enhancements in cancer patient survival while reducing side effects. They remain one of the main areas of focus for many pharmaceutical and biotechnology companies with hundreds of drugs at various stages of the developmental pipeline. Unfortunately, the tumor heterogeneity that ushered in the development of targeted therapies is also what makes targeted therapies similar to chemotherapy and hormonal therapy, the development of drug resistance. The stochastic mutations that cause one tumor to differ from another can also generate two or more subpopulations of cells to arise within a tumor, one sensitive and one resistant to therapy. Hence, it becomes important to understand the breadth of oncogenic mutations within tumors to fully leverage the power of targeted therapies.

1.2 c-MYC oncogene and Mouse Models

The Myc oncogene was discovered in the late 1970's through research on fulminant chicken tumors (53, 54). Since then, it has been recognized as a central transcription factor involved in cell growth, proliferation, tumorigenesis, and even stem cells. Because of its importance, Myc expression is highly regulated at the mRNA levels by numerous mechanisms involving regulatory motifs in its promoter (55, 56) and its protein stability is tightly controlled through the Ras-mediated Raf-MEK-ERK kinase cascade and PI3K-Akt pathway that inhibits GSK-3β (57). In tumorigenesis, it was determined early on that MYC was frequently translocated in Burkitt's lymphoma (58, 59). In addition, Myc is frequently amplified or upregulated in many tumors such as breast, colorectal, pancreatic, gastric, uterine, and ovarian cancers (60, 61). Therefore, Myc is considered one of the most frequent genetic alterations found in human cancer (62).

However, deregulated Myc expression in and of itself is not sufficient for tumorigenesis. Enforced Myc expression actually induces apoptosis through a combination of ARF induction, leading to p53 accumulation and relieving the cell of MDM2-dependent feedback (63), and the suppression of antiapoptotic BCL-2 and BCL-xL (64, 65). Because of this tendency of Myc over-expression to cause cell death, it was hypothesized that the complete malignant transformation would require a cooperating "second hit", another mutation or set of mutations that reduces the apoptotic potential of Myc.

1.2.1 Myc mouse models

Mouse models have contributed greatly to the understanding of Myc as an oncogene. One of the earliest models coupled mouse Myc to the mouse mammary tumor virus promoter (MMTV-Myc) (66). These mice develop spontaneous mammary adenocarcinomas and various iterations of the transgenic line have been used to study cell cycle regulation (67), tumor development (68), and heterogeneity (69). Myc by itself is not sufficient for tumorigenesis, as evidenced by the fact that tumors develop spontaneously, presumably after the second hits, rather than uniformly following Myc induction. Myc serves to drive overgrowth of cells that increases the chances a tumorigenic mutation occurs that allows for tumor development.

Another model couples Myc to the immunoglobulin heavy chain enhancer (Eμ-Myc) (70), mirroring the translocation seen in Burkitt's lymphoma. Mice from the resulting transgenic line develop B-cell lymphomas that arise from both immature and mature B lymphocytes. Similar to MMTV-Myc mice, Myc in this case is also not sufficient for tumorigenesis. Constitutive Myc expression leads to an expanded population of pre-B cells in which a "second hit" may occur, causing tumors to arise at random (71). Clonality experiments have shown that multiple cells can undergo tumorigenesis concurrently, leading to tumors that arise from two or more pre-B and B cells. The second hit has been well studied in the Eμ-Myc model. Mutation analysis of Eμ-Myc tumors has also born out the need for coordinating genetic alterations that decrease the apoptotic potential of Myc overexpression. The majority of arising tumors contain mutations that either disrupt the ARF-MDM2-p53 pathway (63, 72, 73) or cause overexpression of pro-survival BCL-2 or BCL-xL (64, 65, 74). In fact, about 80% of lymphomas that arise from the Eμ-Myc model have an alteration in the p53 axis, whether it be deletion of ARF (24%), mutation or loss of p53 (28%), or Mdm2 overexpression (48%) (63). In this pathway, ARF binds to and sequesters Mdm2 (75, 76), which normally functions to inhibit p53 transactivation (77) and target p53 for degradation by ubiquitination(78, 79). Mdm2 is also a p53 transcriptional target, providing a negative feedback mechanism regulating p53 function (80, 81). BCL-2 and BCL-xL are often overexpressed in a number of human cancers including non-Hodgkin's lymphoma, colorectal, gastric, prostate, and non-small cell lung cancers, and neuroblastomas (reviewed in (82)).

The second-hit model is further supported by results from double transgenic animals. Eμ-Myc mice crossed with mice in which the ARF locus is deleted display much faster lymphomagenesis through suppression of p53 (83). When BCL-2 is also placed under transcriptional control of the immunoglobulin heavy chain enhancer, the resulting transgenic mice when crossed with Eμ-Myc mice also show faster tumor development (74). On the other hand, mice deficient in Mdm2 binding protein (MTBP), which increases p53 degradation through increasing Mdm2 ubiquitin ligase activity, when crossed into the Eμ-Myc model show delayed lymphoma development (84). Taken together, these results indicate the clear need for additional alterations to induce tumorigenesis and that these alterations frequently occur in the p53 axis or BCL2 family.

1.3 The E2F family

The E2F family is a group of transcription factors that are downstream effectors of the tumor suppressor, retinoblastoma (Rb). There are eight genes and nine proteins within the E2F family, which function in many biological processes including DNA replication and repair, mitosis, mitotic and DNA-damage checkpoints, differentiation, development, and apoptosis (85-87). E2F1, E2F2, and E2F3a are considered the activator E2Fs because of their main role in activating gene transcription, while E2F3b and E2F4-E2F8 constitute the repressor E2Fs because of their function in repressing gene expression. Many E2F members have also been closely tied to the oncogenic process.

Rb, an important regulator of the E2F family, was first identified because of its association with a heritable eye tumor, retinoblastoma (88, 89). In non-proliferating cells, Rb is bound to the E2Fs, which suppresses their function (90-96). The binding of Rb to E2Fs also represses transcription through recruitment of histone deacetylases, histone methyltransferases, and DNA methyltransferases to the promoters of E2F target genes (97). Following growth stimulation, Rb is phosphorylated by a complex of cyclin D and cyclin dependent kinase 4 (CDK4) or cyclin D and CDK6 (98). This hypophosphorylated Rb can no longer bind the E2Fs, thus allowing E2F to trigger a transcriptional cascade that leads to progression through the cell cycle. Deregulation of the Rb-E2F pathway, whether through loss of Rb function, cyclin D amplification, or p16/CDKN2A (CDK inhibitor) loss resulting in Rb hyperphosphorylation occurs in the majority of human cancers (99).

Following mitogenic stimulation, activator E2Fs are released to target transcription of many genes involved in cell-cycle progression such as cdc6, an essential DNA replication regulator that has an important role in activating cell cycle checkpoints leading to S phase and mitosis (100), Cyclin E, whose binding to CDK2 in G1 phase is required for transition to S phase and determining cell division (101), and DNA polymerases.

Studies deregulating expression of the activator E2Fs have highlighted their importance in the cell cycle. For example, cells in which E2F1, E2F2, and E2F3 are lost show impaired expression of essential cell-cycle progression genes that are E2F targets and exhibit an inability to progress through the cell cycle (102). Similar to Myc, deregulation of E2Fs are also tied to the p53 axis. Under normal circumstances, ARF is repressed, allowing Mdm2 to bind to and inactivate p53. However, in E2F3 deficient cells, ARF is derepressed (103), leading to p53 activation and defects in the cell cycle including upon re-entry into the cell cycle from quiescence. ARF loss, however, abrogates these defects in E2F3 cells. ARF also activates transcription of E2F transcriptional-repressor complexes, leading to inactivation of activator E2Fs (104). Also, cells without E2F1-E2F3 exhibit p53 activation and interestingly, will express E2F target genes upon loss of p53 (105). Hence, cell proliferation is regulated by E2F through the p53 axis and consequent induction of repressor E2Fs. Deregulation of E2Fs, in and of itself, is often insufficient to cause tumorigenesis and requires a second hit, often in the p53 axis, for cancer development.

1.3.1 E2F1

Among the activator E2Fs, E2F1 is unique in its ability to directly induce apoptosis, as seen when E2F1 is ectopically expressed (106). This is further evident in the fact that mice deficient in E2F1 will develop tumors in part because of the lack of E2F1-mediated apoptosis (107). E2F1 can induce apoptosis both through p53 and also independently of p53.

In p53-dependent apoptosis, abnormal E2F1 expression can stabilize and activate p53 to cause apoptosis by multiple mechanisms. First, E2F1 may activate ARF expression that leads to p53 stabilization and activation (108). Additionally, E2F1 can also induce p53 phosphorylation at certain amino acids that is normally seen in response to DNA damage (109, 110), leading to p53 activation in an ARF-independent manner. E2F1 can also upregulate p53 co-factor (ASPP1 and ASPP2) expression that directs p53 towards activating pro-apoptotic genes and subsequent apoptosis (111-113). Rb is also a participant in E2F1-mediated apoptosis through p53, as Rb inactivation will induce p53-mediated apoptosis. This can be inhibited by loss of E2F1 (114). This link between the Rb-E2F pathway and the p53 pathway can be seen in the fact that because Rb inactivation leads to deregulated E2F1 and subsequent E2F1 mediated apoptosis, a cooperating mutation in the p53 axis often occurs with Rb loss in tumors (99, 115).

E2F1-induced apoptosis that is independent of p53 occurs through a variety of pathways. E2F1 activates expression of APAF1 (116), which binds to cytochrome c and ATP to form an apoptosome that cleaves and activates caspase 9 to stimulate the caspase cascade, committing the cell to apoptosis (117, 118). E2F1 also upregulates other caspases as well as p73 (116), structurally related to p53 and involved in apoptosis (119). It also directly upregulates Bcl-2 homology 3 (BH3)-only genes, which are proapoptotic Bcl-2 family members, such as PUMA, Noxa, and Bim (120). Along with upregulation of pro-apoptotic genes, E2F1 also represses expression of survival signals mediated by NF-κB and BCL-2 (116).

Given E2F1's potential to induce both proliferation and apoptosis and its importance in initiating the cell cycle, the apoptotic arm of E2F1 must be tightly regulated. Rb has a domain called the Rb pocket that binds E2Fs. It has an additional binding site that binds E2F1, which specifically inhibits E2F1-mediated apoptosis (121). Following DNA damage, this region of Rb becomes acetylated, releasing E2F1 without affecting Rb binding to other E2Fs (122). E2F1 also becomes acetylated following DNA damage and these changes favor the induction of pro-apoptotic E2F1 targets such as p73 (123). Moreover, cells deficient in E2F1 show reduced apoptosis following DNA damaging chemotherapy treatment (124, 125). These reports indicate that E2F1-mediated apoptosis is a response of DNA damage. E2F1 binding partners also mediate E2F1-induced apoptosis. Jab1 binds to the E2F1 marked box domain, one that has proapoptotic activity unique to E2F1 (126), and its interaction with E2F1 directs expression of proapoptotic genes (127). Similarly, Foxo1 and Foxo3 also bind E2F1 to direct a proapoptotic transcriptional program that includes APAF1 and CDKN1C (128). API5, an apoptosis inhibitor, is another mediator of E2F1-induced apoptosis but does not serve to inhibit E2F1-mediated transcription (129).

Pathways are also involved in the regulation of E2F1-induced apoptosis. The PI3K-Akt pathway in particular, specifically inhibits this activity. When PI3K, a family of signal transducing kinases, is activated by growth factors such as IGF1, it activates Akt, serine/threonine kinase (130). Akt affects E2F1-induced apoptosis by phosphorylating TopBP1, allowing it to bind E2F1 and repress E2F1-induced apoptosis (131). The PI3K-Akt pathway also inhibits E2F1-mediated apoptosis by repressing E2F1 targets involved in the process (132). Breast and ovarian cancers that exhibit low expression of this specific set of genes show worse prognoses. Summarily, E2F1-mediated apoptosis is a tightly controlled process normally induced by DNA damage or expression of E2F1 without growth signals. When the process of E2F1-induced apoptosis is deregulated through mutations that affect any of the genes or pathways mentioned above, cells are subsequently desensitized to apoptosis, providing an opportunity for deregulated growth and tumorigenesis.

2. Genomic analysis of the Eμ-Myc mouse as a model for heterogeneity in therapeutic response

2.1 Introduction

Aggressive B-cell lymphomas include a spectrum of diagnoses that span Burkitt lymphoma (BL), diffuse large B-cell lymphoma (DLBCL), and lymphomas that lie between these diagnoses, termed by the World Health Organization 2008 classification as "B-cell lymphoma, unclassifiable, with features intermediate between DLBCL and BL"(133). There is considerable clinical and therapy response heterogeneity across and within these diseases. While diffuse large B-cell lymphoma is generally responsive to the R-CHOP chemo-immunotherapy regimen (containing rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone) that is commonly used for treatment of non-Hodgkin's lymphoma, Burkitt's lymphoma requires more aggressive multi-agent regimens that are accompanied by higher toxicities. Although these lymphoma subtypes are generally treated differently, patients are not always cured and responses are not always complete. Prior studies evaluating the heterogeneity of aggressive B-cell lymphomas using primary patient samples have begun to highlight that biologic and genomic complexity underlies clinical variation (10, 134-137). More recently, exome, RNA, and whole genome sequencing have highlighted the recurrent and presumably driver mutations that frequently occur in BL and DLBCL and that many mutations are specific to BL or a subtype of DLBCL (138-142). Currently, these studies have pointed to possible biomarkers to aid in therapy selection. None have yet made it to clinical use, however. The inherent limitations in the availability and quality of patient-derived samples suggest that experimental models could greatly facilitate efforts to understand heterogeneity in aggressive lymphomas and in therapy response and to develop appropriate therapeutic options.

Genetically engineered mouse models have provided significant insight into human cancer biology. These models result from the activation or loss of a single gene, and are generally considered to represent a distinct and homogenous phenotype. However, our previous work has provided evidence of heterogeneity in genetically engineered mouse models, specifically the MMTV-*Myc* model of breast cancer and the *Eμ-Myc* model of B-cell lymphoma (69, 143). In both cases, large numbers of tumors from these transgenic mice were evaluated and found to have natural heterogeneity in histologic characteristics and genome-scale expression data exists, suggesting secondary genetic hits drive variation in genetically engineered mouse models.

The *Eμ-Myc* transgenic mouse was developed as a model of *Myc*-driven aggressive B-cell lymphoma (70, 71). Multiple investigations have identified genes that alter the onset of *Eμ-Myc* lymphoma and/or affect response to single agent chemotherapy (74, 144-151). The *Eμ-Myc* model has also been used in a genetic screen to identify genes that modulate response to doxorubicin (152). By focusing largely on perturbing single genes in the *Eμ-Myc* background, these studies did not focus on the effects of gene networks and their relevance to the natural genetic heterogeneity seen in the *Eμ-Myc* model, particularly as it pertains to response to lymphoma therapy.

While the *Eμ-Myc* model does show similarities to human disease, there are a number of limitations. Cell of origin is one such shortcoming. The majority of *Eμ-Myc* tumors tend to be pre-B in surface marker expression, owing to the transgene causing clonal expansion of pre-B cells, with a minority being mature B cells (71). DLBCL, on the other hand, is predominantly derived from mature B cells, whether they be germinal center of activated B cell, and BL are typically germinal center B cells. Furthermore, exome, RNA, and whole genome sequencing have shown recurrent gene mutations, some of which are enriched in BL or a subtype of DLBCL, such as ID3 in BL, EZH2 in the germinal center B-cell like subgroup of DLBCL, and MYD88 in the activated B-cell like subgroup of DLBCL (138-142). Lack high-throughput sequencing in *Eμ-Myc* tumors has prevented a comparison with the complexity of the mutational landscape in human lymphoma.

We previously described that the *Eμ-Myc* transgenic mouse model of lymphoma develops genomically distinct lymphoma subtypes that reflect the spectrum of human aggressive B-cell lymphomas (143). Here, we describe a genomic analysis strategy to reproducibly classify the distinct forms of *Eμ-Myc* tumors and methods to use the *Eμ-Myc* lymphoma model to predict therapy response.

2.2 Materials and Methods

2.2.1 Mouse strains and tumor monitoring

Eμ-Myc strain (JAX stock# 002728) and C57BL/6J strain (JAX stock# 000664) mice were purchased from Jackson Laboratories, and were bred and housed in a Duke University Medical Center Division of Laboratory Animal Resources facility. All experiments were approved by the Duke University Institutional Animal Care and Use Committee. *Eμ-Myc* mice were monitored twice weekly for visible or palpable lumps, a hunched posture, tachypnea, a swollen belly, or ruffled fur. Upon development of such symptoms, mice were sacrificed and dissected. The spleen and enlarged lymph nodes or masses were removed and placed in RPMI media (GIBCO) with 10% heat-inactivated fetal bovine serum (FBS, Sigma). Spleens were weighed and lymph node specimens were frozen in liquid nitrogen or fixed in 10% formalin. Single cell suspensions were generated from remaining tissue by squeezing between ground glass slides and filtering the suspension through a sterile 100 μm cell strainer (BD Falcon). Thereafter, we lysed red cells in the cell suspensions and washed the cells twice. Lymphoma cells were either stored inviably at -80°C as cell pellets or resuspended in freezing media (10% DMSO (Sigma) in heat-inactivated FBS), aliquoted, and stored viably in liquid nitrogen.

2.2.2 Generation of *Eμ-Myc* lymphoma in C57BL/6 mice and treatment

To transplant *Eμ-Myc* lymphomas to C57BL/6 mice, we thawed and washed frozen *Eμ-Myc* lymphoma cells and determined the viable cells. We suspended the lymphoma cells in RPMI media (GIBCO) and injected 5x10^5 lymphoma cells by the *i.p.* route into C57BL/6J mice. We monitored recipient mice daily, and rated illness on a scale from 1 to 5, where 1 was moribund, 3 was obvious lymphoma, and 5 was perfect health. Mice with a rating of 3 were either sacrificed or treated with chemotherapy – cyclophosphamide (300 mg/kg *i.p.* once), doxorubicin (10 mg/kg *i.p.* once), or bortezomib (1 mg/kg *i.p.* twice weekly for four doses) – and were thereafter monitored daily using the rating scale until progression (score less than 3), at which time they were sacrificed and lymphoma tissue was collected. Cyclophosphamide and doxorubicin were purchased from the Duke University Medical Center inpatient pharmacy, and bortezomib was purchased from LC laboratories.

2.2.3 Gene expression profiling data preparation and normalization

Lymphoma tissue samples were homogenized using the Lysing Matrix A tube for tissue (MP Biomedicals). Lymphoma cell pellet samples were homogenized by passing through a tuberculin syringe and centrifuged through a QIAshredder homogenizer column (Qiagen). RNA was extracted from the lysates using Qiagen RNeasy kits (Qiagen). RNA integrity was verified with an Agilent 2100 Bioanalyzer. Microarray processing and RNA hybridization to Affymetrix Mouse 430 2.0 or Mouse 430A 2.0 GeneChip arrays were performed according to the manufacturer’s instructions in the Duke University DNA microarray core facility.

From microarray CEL files, data was normalized with either MAS5 or RMA algorithms for appropriate downstream analyses. Gene expression data can be accessed in the NCBI Gene Expression Omnibus (GSE40760). For analyses comparing data obtained from microarrays processed at different times (separate batches), the ComBat algorithm (153) was used to reduce batch effect. For analyses comparing human and mouse microarray data, we used ChipComparer[1] and FileMerger[2], as described previously (69).

2.2.4 Microarray analysis

We performed unsupervised analyses of unfiltered genomic data from *Eμ-Myc* lymphomas using the k-means clustering algorithm. For supervised analyses, we determined which genes were differentially expressed between the two clusters using the Wilcoxon rank sum test and Bonferonni correction (q value < 0.2). Gene ontology and biological function of differentially expressed genes was assessed using DAVID[3], and significant gene ontology terms were selected from Bonferonni values less than 10^{-5}. We also used genomic data to develop new genomic models ("signatures") from existing microarray datasets using the binary regression (Binreg) algorithm (for two-class models) (20, 154) or Prediction Analysis for Microarrays (PAM; for greater than two-class models (155). Datasets used to generate and validate signatures along with

---

[1] http://chipcomparer.genome.duke.edu/

[2] http://filemerger.genome.duke.edu/

[3] http://david.abcc.ncifcrf.gov/home.jsp

accuracy of the signature using leave-one-out cross validation are listed in Table 1. Parameters used in generating genomic signatures are listed in Table 2. Genomic signatures of cellular pathway activation were applied using the ScoreSignature module (156) on the Duke University GenePattern server.

Table 1: Datasets used for genomic analyses

Database	Dataset Number	Signature	Training Set or Validation Set	Accuracy by Leave-One-Out Cross Validation
GEO	GSE7897	*Eμ-Myc* Cluster	Training	97%
GEO	GSE4475	GCB-ABC; CHOP resistance	Training	100%, 84%
GEO	GSE10846	GCB-ABC; CHOP resistance	Validation	Not Applicable
GEO	GSE26408	B-cell maturation	Training	96%

Table 2: Parameters used for genomic signature generation and application with BinReg algorithm

Signature	Genes (N)	Factors (N)	Burn-in	Iterations	Quantile Normalize?	Standardize?
Eμ-Myc Cluster	500	2	1000	5000	Y	Y
GCB-ABC	200	2	1000	5000	Y	Y
CHOP Resistance	200	2	1000	5000	Y	Y

2.2.5 Immunoblot analysis

Protein was extracted from frozen lymphoma tissue and immunoblots were performed as described previously (147). 100 μg protein samples were fractionated by gel electrophoresis and transferred to Immobilon membrane to detect c-MYC, p19^{ARF}, MDM2, p53, p65/RELA, phosphorylated p65/RELA, β-Actin, and GAPDH. Antibodies used were as follows: c-MYC (Santa Cruz N-262 sc 764, 1:500), p19^{ARF} (Calbiochem Ab-1 PC435, 1:1000), MDM2 (Santa Cruz c-18 sc812, 1:1000), p53 (Calbiochem Ab-1 OP03, 1:1000), p65/RELA (Santa Cruz C-20 sc-372, 1:1000), phosphorylated p65/RELA (Cell Signaling 93HI 3033S, 1:1000), β-Actin (Cell Signaling 13E5 4970S, 1:5000), and GAPDH (Santa Cruz FL-335 sc-25778, 1:500). Equal protein loading was verified by staining blots with Ponceau Red and probing for GAPDH or β-Actin.

2.2.6 Southern analysis

Genomic DNA was isolated from frozen lymphoma tissue, digested with *Eco*R1 (10 mg per sample) and fractionated by agarose gel electrophoresis. Thereafter, DNA was transferred to membranes and probed with a radiolabeled heavy chain J3-J4 joining region genomic fragment, as described before (147).

2.2.7 Flow cytometry

Eμ-Myc lymphoma surface expression of B220, IgM, IgD, CD43, and CD138 were assessed with flow cytometry, using the following reagents: Mouse BD Fc block (BD Pharmingen 553142) and antibodies (B220-APC BD Pharmingen 553092, IgM-FITC BD Pharmingen 553437, IgD-PE BD Pharmingen 558597, CD43-FITC BD Pharmingen 553270, and CD138-PE BD Pharmingen 553714). Cells were washed, incubated with block then antibody, washed again, and fixed prior to assessment on a Becton Dickenson

FACSCanto II flow cytometer. Data were analyzed with FlowJo (Tree Star, Inc.). We also stained and performed flow cytometry on pooled bone marrow, mesenteric lymph node and spleen cells from healthy C57BL/6 mice, which were used as a normalization control between experiments.

2.2.8 Statistical analyses

Statistical analyses were performed with the statistical environment R using standard packages and the pamr and survival packages. Binary regression modeling was performed using MATLAB (21). Pathway signature scores were generated using the ScoreSignature module in GenePattern (156). Progression free survival was calculated as days from therapy to when the mouse had consistent progression of illness compared to the day of therapy, was moribund and sacrificed, or died.

2.2.9 *Trp53* sequence analysis

RNA was isolated from frozen lymphoma cell pellets: cells were lysed in Qiagen RLT buffer, the lysate homogenized by passage through a 20 gauge needle followed by a QIAshredder spin column, and the RNA purified using an RNeasy spin column (Qiagen). The preparation was treated with DNase I (New England Biolabs) and then further purified using a second RNeasy spin column. The RNA was analyzed for quality using the RNA 6000 Pico Kit and the Agilent 2100 Bioanalyzer (Agilent Technologies). RNA (1 μg) was reverse-transcribed using the qScript cDNA SuperMix (Quanta BioSciences). The coding region of the *Trp53* cDNA was PCR amplified in three overlapping segments using the Phusion High Fidelity PCR Master Mix with HF Buffer (New England Biolabs). The size and quantity of the PCR products were verified by agarose electrophoresis. The PCR products were treated with ExoSAP-IT (Affimetrix) and Sanger sequenced. FinchTV was used for the chromatogram analysis. Sequences were analyzed for mismatches and insertions or deletions using BioEdit and the Basic Local Alignment Search Tool (BLAST) hosted by the NCBI.

2.3 Results

2.3.1 A method to predict *Eμ-Myc* lymphoma subtypes

We previously described the intrinsic heterogeneity of global gene expression in lymphomas that develop from the *Eμ-Myc* mouse model (143). Our current unsupervised analysis of gene expression profiling data from 112 *Eμ-Myc* lymphomas confirms variation in overall genomic expression (Figure 1A), and identifies a natural division of these lymphomas into two *Eμ-Myc* subgroups, or “clusters” (Figure 1B). Iterative unsupervised k-means clustering analyses demonstrate reproducibility (99.9% of runs), whereas attempts to subdivide the dataset into more than two clusters resulted in instability in the cluster assignments for the samples.

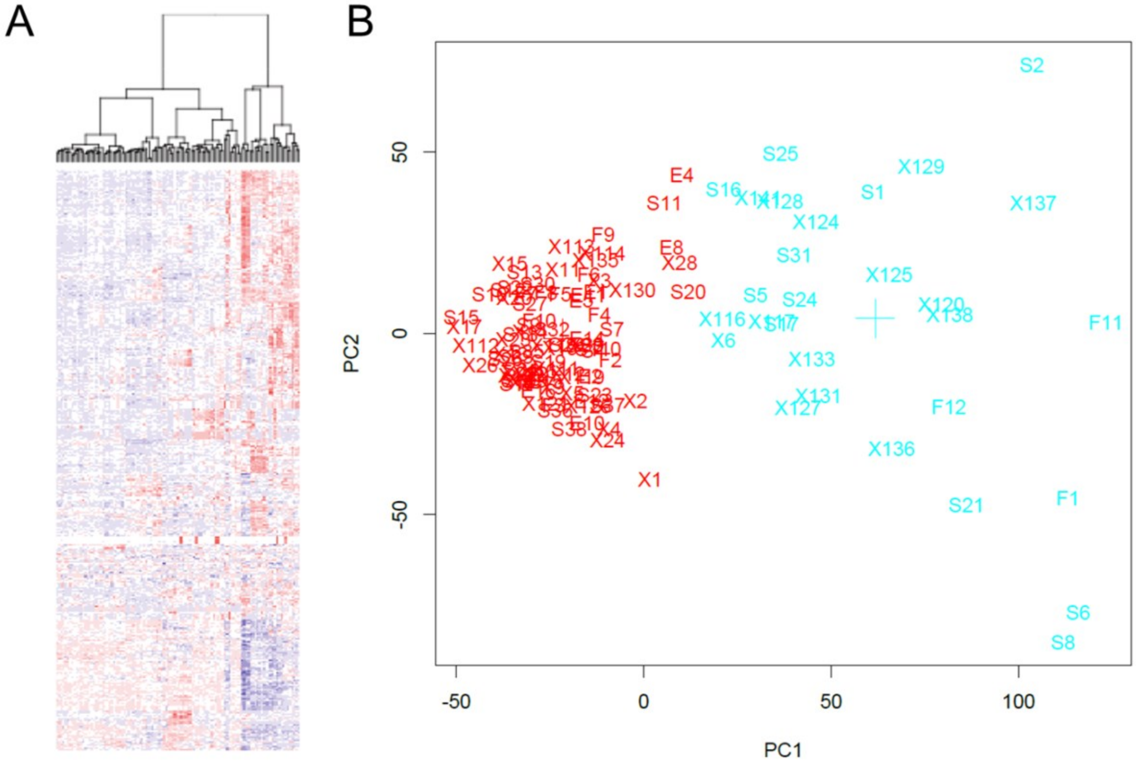


Figure 1. Unsupervised clustering of *Eμ-Myc* lymphomas

A) Heatmap of the *Eμ-Myc* lymphoma training set, with lymphomas clustered based on similarities in expression of the 2% of probes with the highest standard deviation across all samples. Data was centered and scaled; red color represents upregulated gene probes and blue color represents downregulated gene probes. B) The first and second principal component (PC) of gene expression array data from each *Eμ-Myc* lymphoma in the training set are plotted, grouped using k-means clustering by color into Cluster 1 (red) and Cluster 2 (blue). Plus sign represents the k-means clustering centroid. This demonstrates two distinct genomic clusters defined by an unsupervised clustering method.

Moving beyond descriptive analyses, we developed a prediction model (or "signature") that can be used to classify a new lymphoma sample into one of the two clusters (Figure 2A). This 500-probe cluster classification signature has 97% accuracy of correctly classifying training data on leave-one-out cross validation. We evaluated a test set of 76 new *Eμ-Myc* lymphomas, originating in C57BL/6 congenic mice, in which the genomic signature identified 56 Cluster 1 lymphomas and 20 Cluster 2 lymphomas, using a cut-off score of 0.5. Concordant with the prior description (143), there is a significant difference in time to onset of these two types of lymphoma, with Cluster 1 lymphomas more likely to occur at an earlier age than Cluster 2 lymphomas (median time to onset of 121 vs. 326 days, respectively, p = 0.0003, log-rank test, Figure 2B).

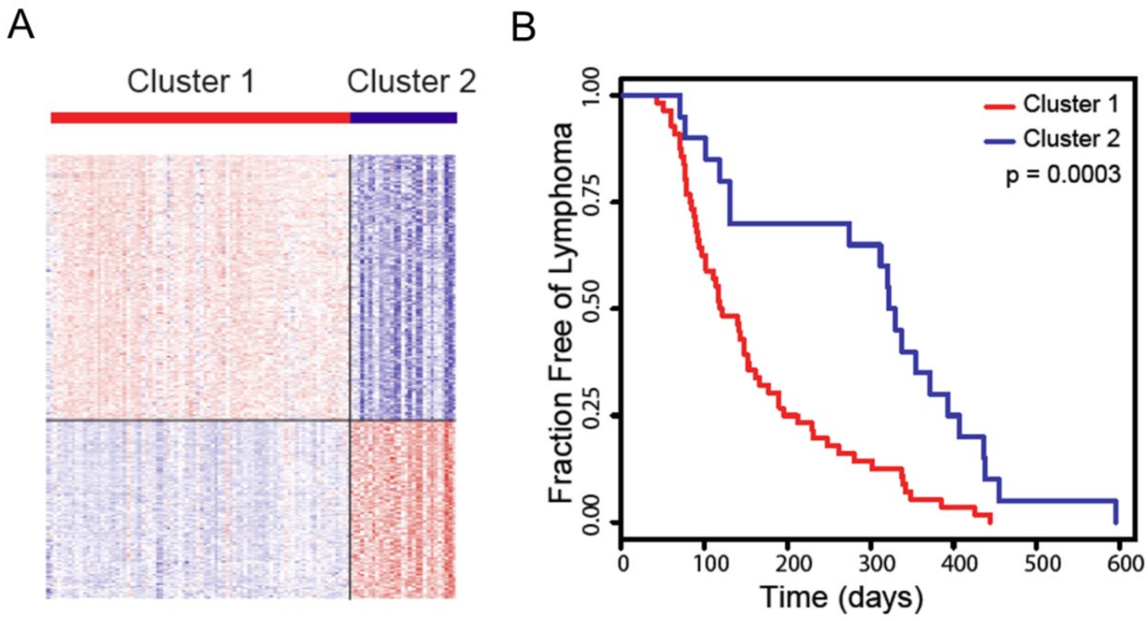


Figure 2. Two clusters of *Eμ-Myc* lymphomas with differential survival

A) Heatmap of the 500-probe genomic signature that discriminates Cluster 1 and Cluster 2 *Eμ-Myc* lymphomas, with samples in columns and probes in rows. Color scheme as in Figure 1A. B) Kaplan-Meier curves evaluating time from birth to onset of lymphoma for *Eμ-Myc* mice, grouped by genomic lymphoma cluster (Cluster 1: n=56, Cluster 2: n=20). Significance determined by the log-rank test.

These results confirm heterogeneity in the transgenic *Eμ-Myc* model and establish a predictive framework that can prospectively evaluate new and independent *Eμ-Myc* lymphomas.

2.3.2 Genetic, biologic, and clinical differences between the *Eμ-Myc* subtypes

Moving beyond simply classifying *Eμ-Myc* lymphomas, we evaluated differences between the identified genomic clusters. For example, in supervised analyses of differentially expressed genes between the two clusters, we found that genes upregulated in Cluster 1 lymphomas are significantly enriched for gene ontology biological function terms related to RNA processing, regulation of transcription and translation, and cell cycle, whereas genes upregulated in Cluster 2 lymphomas are significantly enriched for gene ontology biological function terms that include immune response, protein localization, and regulation of apoptosis (p values all $< 10^{-8}$; Bonferroni FWER all $< 10^{-5}$, Table 3).

Table 3: Highest Gene Ontology Biological Processes Terms that are significantly upregulated in Cluster 1 vs. Cluster 2 *Eμ-Myc* lymphomas

Gene Ontology Term	Upregulated in *Eμ-Myc* Lymphoma Cluster	Nominal P-value	Bonferonni value
RNA processing	Cluster 1	1.92E-111	5.99E-108
mRNA metabolic process	Cluster 1	3.92E-76	1.22E-72
mRNA processing	Cluster 1	1.97E-75	6.14E-72
RNA splicing	Cluster 1	1.47E-69	4.60E-66
Cell cycle	Cluster 1	1.23E-59	3.83E-56
DNA metabolic process	Cluster 1	1.99E-57	6.22E-54
Immune response	Cluster 2	4.68E-20	1.46E-16
Protein localization	Cluster 2	2.06E-11	6.42E-08
Establishment of protein localization	Cluster 2	8.24E-11	2.56E-07
Protein transport	Cluster 2	1.14E-10	3.56E-07
Vesicle-mediated transport	Cluster 2	2.54E-10	7.90E-07
Regulation of apoptosis	Cluster 2	4.08E-09	1.27E-05

A further evaluation of gene expression data identified significant differences between the *Eμ-Myc* lymphoma clusters in terms of gene sets that represent B-cell maturation. Using a stage-specific genomic signature developed from sorted murine B-cells of different maturation stages (157), we classified *Eμ-Myc* lymphomas as pro/pre B-cell stage (n = 53, 70%), follicular/marginal zone stage (n = 14, 18%) and germinal center B-cell stage (n = 9, 12%). Lymphomas classified as deriving from pro/pre B-cells were more common in the Cluster 1 subgroup, whereas lymphomas classified as from the follicular/marginal zone stage were more common in Cluster 2 subgroup ($p = 5x10^{-7}$, Chi-squared test, Table 4).

Table 4: Distribution of *Eμ-Myc* lymphomas with regard to B-cell maturation stage, determined by genomic analyses

	Cluster 1 (n = 56)	Cluster 2 (n = 20)
Pro/Pre B-cell	48	5
Follicular/Marginal Zone B-cell	3	11
Germinal Center B-cell	5	4

Since stage of differentiation and activation has prognostic value in human DLBCL (10), we evaluated the *Eμ-Myc* lymphoma genomic data within the context of germinal center B-cell (GCB) versus activated B-cell (ABC) lymphoma subtypes. While there is variation in the stage of differentiation of DLBCL, BL arise uniformly from B-cells at the GCB stage (134, 158). Using genomic data from human lymphoma samples (159), we evaluated a GCB/ABC genomic signature in human and *Eμ-Myc* lymphomas (Figure 3A). In an independent dataset of human DLBCL samples (135), the signature correctly distinguished lymphomas annotated as GCB or ABC type (Figure 3B, $p < 2x10^{-16}$, Wilcoxon rank-sum test), and when a cut-off score of 0.5 was used, the accuracy was 94% (n=329/350). Applying this signature to *Eμ-Myc* lymphoma data identified significant differences between the two cluster subtypes (Figure 3C, respective median scores 0.18 vs. 0.99, where 0 represents GCB and 1 represents ABC, $p < 0.0001$, Wilcoxon rank-sum test). Using a cut-off score of 0.5, Cluster 1 lymphomas were more likely to have a GCB score, while Cluster 2 lymphomas were more likely to have an ABC score ($p = 0.0001$, Pearson's Chi-squared test). These different supervised analyses of genomic data underscore the distinct differences between *Eμ-Myc* lymphomas and suggest there is variation in cell of origin maturation and differentiation.

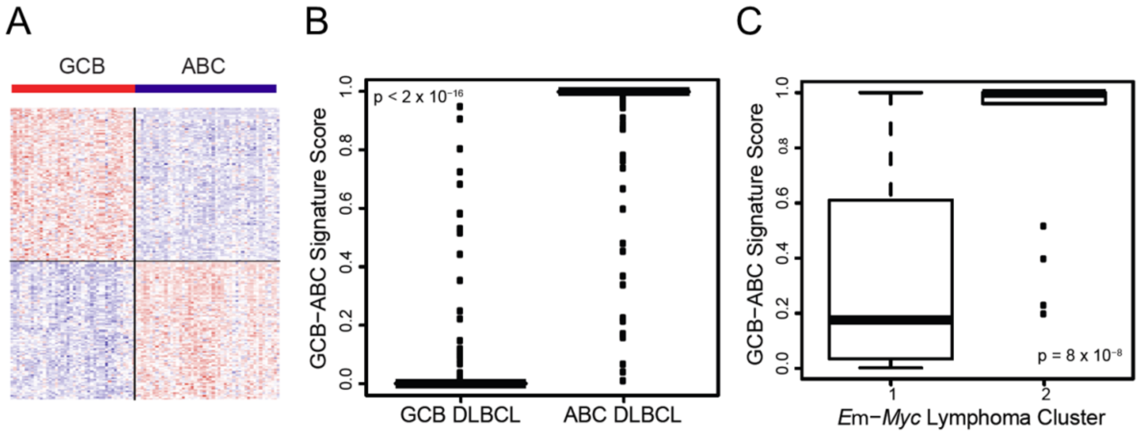


Figure 3. GCB/ABC expression patterns segregate in *Eμ-Myc* lymphomas

A) Heatmap of the genomic signature that discriminates germinal center B-cell (GCB) from activated B-cell (ABC) lymphomas. Red color represents upregulated gene probes and blue color represents downregulated gene probes. B) Box and whisker plot of GCB-ABC signature scores in an independent dataset of human B-cell lymphomas, grouped by GCB or ABC lymphoma type. Signature score of 0 represents GCB, while 1 represents ABC. Bold line represents median, box represents the interquartile range (IQR), whiskers represent 1.5 times the IQR, and dots represent outliers. P value calculated using the Wilcoxon rank sum test. C) Box and whisker plot of GCB-ABC signature scores in Cluster 1 and Cluster 2 *Eμ-Myc* lymphomas. Graphing parameters as described in Figure 3B. p-value calculated by Wilcoxon rank sum test.

Our prior work linked *Eμ-Myc* lymphoma subgroups with human aggressive lymphoma subtypes. We evaluated genomic data from aggressive human lymphomas (the BL to DLBCL spectrum) and confirmed that the genomic signature we developed to classify *Eμ-Myc* lymphomas significantly differentiates molecularly defined human aggressive B-cell lymphoma subtypes (where a score of zero represents Cluster 1 and a score of one represents Cluster 2). BL is most similar to Cluster 1 *Eμ-Myc* lymphomas, while DLBCL (non-molecular BL) is most similar to Cluster 2 *Eμ-Myc* lymphomas (Figure 4). Together with the analysis of genomic data with regards to GCB versus ABC distinction, these results define Cluster 1 *Eμ-Myc* lymphoma as a representation of human BL and Cluster 2 *Eμ-Myc* lymphoma as a representation of the ABC subtype of human DLBCL.

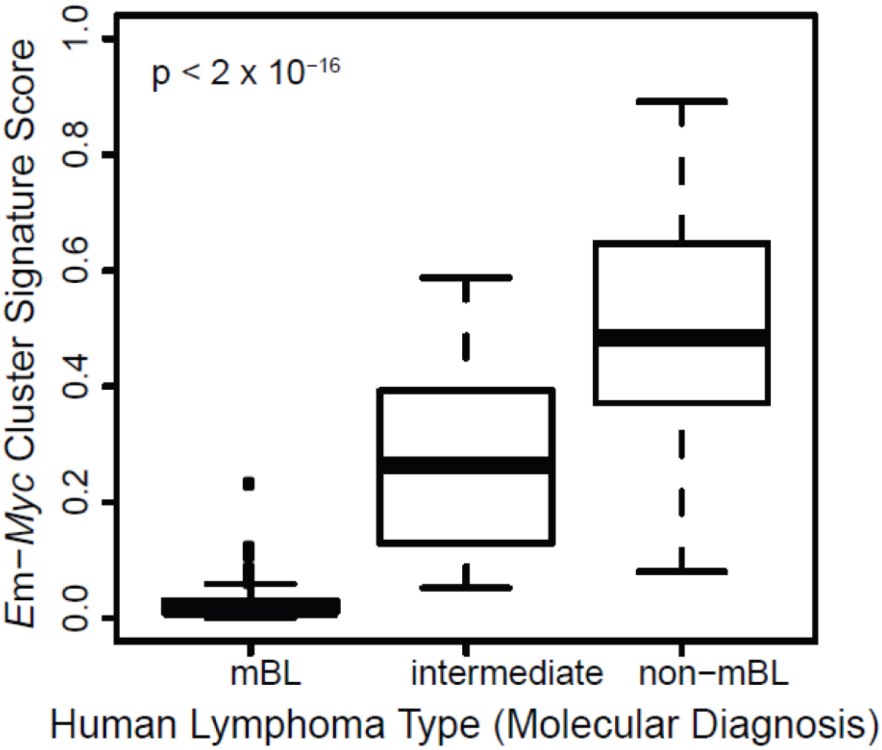


Figure 4. ***Eμ-Myc*** **signature segregates human aggressive disease**

Box and whisker plot of *Eμ-Myc* signature scores in three types of aggressive human B-cell lymphomas (data from GSE4475 dataset). Bold line represents median, box represents the interquartile range (IQR), whiskers represent 1.5 times the IQR, and dots represent outliers. P value calculated using the Kruskal-Wallis test.

In an additional evaluation of complexity in the Eμ-*myc* model, we performed array CGH on DNA from 74 Eμ-*myc* lymphomas (50 Cluster 1 and 24 Cluster 2). We found significant differences between Cluster 1 and Cluster 2 lymphoma in alterations at chromosomes 3, 4, 7, 14, and 15 (Figure 5). For example, one region on chromosome 14 was significantly deleted in Cluster 2 lymphomas, while another region on chromosome 15 was significantly amplified in Cluster 2 lymphomas. Notably, many genes identified within these regions are altered in human B-cell lymphomas. For example, the deleted region of chromosome 14 in Cluster 2 Eμ-*myc* lymphomas contained genes known to be deleted or mutated in DLBCL, including Tnfrsf10b and Dock5 (160, 161). Additionally, the amplified region of chromosome 15 in Cluster 2 Eμ-*myc* lymphomas contained genes known to be mutated and/or constitutively activated in DLBCL, such as Irak4, Skp2, and Pim3 (162-164). Thus, recurrent chromosomal alterations underscore the heterogeneity observed in the Eμ-*myc* model and identify common genetic alterations between murine and human lymphomas.

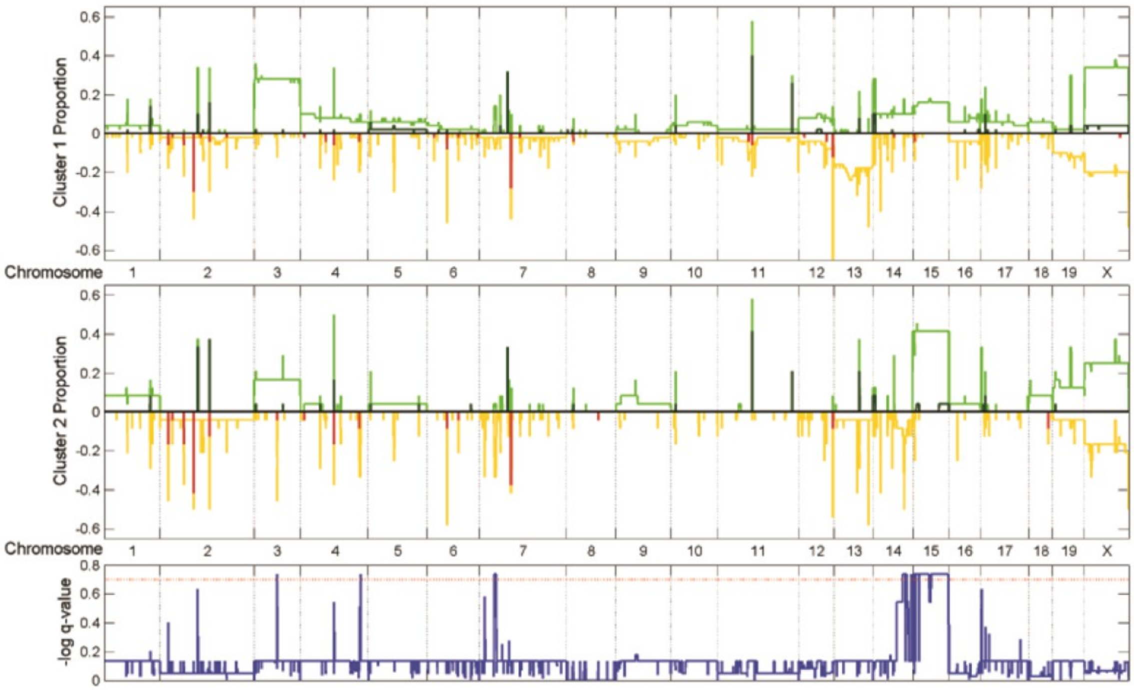


Figure 5. Array CGH of *Eμ-Myc* lymphomas

Top: The proportion of lymphomas in Cluster 1 (top) and Cluster 2 (middle) with copy number aberration, in order by genes on the chromosomes. High copy amplifications denote 2 or more extra copies at a specific locus. Bottom: Analysis of significant copy number differences between Cluster 1 and Cluster 2 lymphomas. A cut-off q-value (determined by false-discovery rate) of 0.2 was used to determine genes with significant copy number changes (bottom). This cut-off is represented by the dotted red line. These results identify chromosomal regions with significantly different rates of amplification or deletion between the two transcriptionally-defined lymphoma clusters.

In addition to genomic evaluations of Eμ-*myc* lymphomas, we assessed surface expression of B-cell markers by flow cytometry. While individual lymphomas varied in expression of IgM, IgD, CD43, and CD138, we found no significant difference between the Eμ-*myc* lymphoma clusters in terms of surface marker expression.

We also evaluated the mutational status of the *Trp53* gene and the overall status of the p53-p19^{ARF}-MDM2 tumor suppressor axis in representative clones of the two *Eμ-Myc* clusters, as past work has documented mutation of p53 in over a quarter of *Eμ-Myc* lymphomas and overall disruption of the p53-p19^{ARF}-MDM2 tumor suppressor axis in about 80% of lymphomas arising in this mouse model (63). Five out of seven Cluster 1 lymphomas harbored mutation in the *Trp53* cDNA, including extensive deletions and others that would lead to missense amino acid substitutions or frameshifts, while all seven of the Cluster 2 lymphomas evaluated contained wild type *Trp53* cDNA ($p = 0.026$, Chi-squared test) (Table 5). In addition, one of the Cluster 1 lymphomas with wild type p53 exhibited elevation of MDM2 expression, whereas the other had elevated p19^{ARF} expression. Five of the Cluster 2 lymphomas exhibited moderate elevation of p19^{ARF} expression, two had moderate increase in MDM2 expression, and one had substantial increase in MDM2 expression (Figure 6). As such, it appears that *Trp53* gene disruption is an additional distinction between the two lymphoma clusters. However, given the observed heterogeneity within and between the lymphoma clusters with regards to mutation of *Trp53* and expression of p19^{ARF} and MDM2, *Trp53* mutation status and the p53 tumor suppressor axis do not appear to be the sole determinants of the differences between the clusters.

Table 5: *Trp53* mutation status

Lymphoma Clone	***Trp53* status**
C1-1	L124P
C1-2	Wild type
C1-3	Y120R with 2 nucleotide insertion
C1-4	Deletion of exons 2 through 7
C1-5	Wild type
C1-6	R242C
C1-7	No product – presumed deletion of locus
C2-1 through C2-7	Wild type

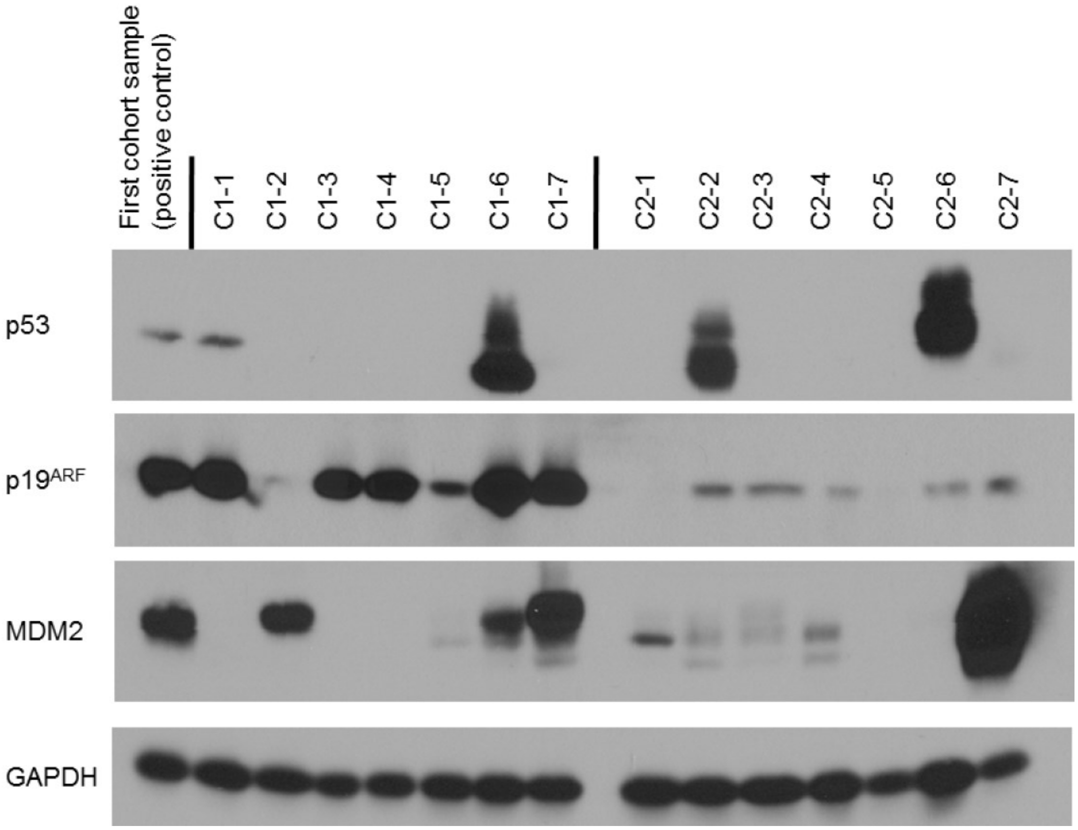


Figure 6. p53 axis protein expression

Immunoblots for p53, p19^{ARF}, and MDM2 protein expression in fourteen spontaneously arising *Eμ-Myc* lymphomas from the new cohort (seven Cluster 1 and seven Cluster 2). A sample from the original *Eμ-Myc* cohort served as a positive control. GAPDH expression demonstrates equal protein loading.

In addition to genomic evaluations of *Eμ-Myc* lymphomas, we assessed surface expression of B-cell markers by flow cytometry. While individual lymphomas varied in expression of IgM, IgD, CD43, and CD138, we found no significant difference between the *Eμ-Myc* lymphoma clusters in terms of surface marker expression (data not shown).

From a clinical perspective, *Eμ-Myc* lymphomas are classically described as presenting with diffuse lymphadenopathy (70). We observed this presentation in the majority of *Eμ-Myc* lymphomas, but also found sick mice with a single enlarged lymph node, mediastinal disease, or gut-centered disease. Cluster 1 lymphomas were significantly more likely to present with lymphadenopathy that was diffuse or involving more than one lymph node region, whereas Cluster 2 lymphomas were significantly more likely to present with lymphadenopathy isolated to one lymph node group, particularly gut-centered (p = 0.0005, Pearson's Chi-squared test, Table 6).

Table 6: Distribution of *Eμ-Myc* lymphomas with regard to phenotype

	Cluster 1 (n = 56)	Cluster 2 (n = 20)
Diffuse lymphadenopathy	45	6
Mediastinal, isolated	7	3
Gut infiltration, isolated	0	8
Mediastinal and gut	4	3

Taken together, our evaluations of *Eμ-Myc* lymphomas categorized into two clusters based on gene expression patterns not only describe biological and clinical differences between the clusters but also link them to distinct human lymphoma entities. Thus, genomic heterogeneity within the transgenic *Eμ-Myc* mouse model appears to mirror genomic heterogeneity in human lymphomas.

2.3.3 *Eμ-Myc* lymphomas exhibit variation in chemotherapy response that reflects variation in chemotherapy response of human aggressive lymphomas

Human aggressive B-cell lymphomas display heterogeneity with regards to response to therapy. Within DLBCL, clinical markers (IPI score (165, 166)) or molecular markers (ABC vs. GCB type (10)) distinguish patients with variable responses to multi-agent chemotherapy or chemo-immunotherapy. BL is uniformly treated with high-dose multi-agent regimens (167), and retrospective analyses revealed that DLBCLs that are genomically similar to BL have improved outcomes when treated intensively like BL (134).

Given the links between human aggressive B-cell lymphomas and the *Eμ-Myc* model, we were interested in determining the extent to which the *Eμ-Myc* model could be used to model heterogeneity in response to lymphoma therapy. We used gene expression signatures as a method by which to generate hypotheses regarding therapy response, and followed these analyses with experimental testing using *Eμ-Myc* tumors.

We began by utilizing a genomic signature developed to distinguish lymphomas that are sensitive or resistant to CHOP-like chemotherapy regimens (Figure 7A) (159). Interestingly, *MYC* was differentially expressed in the 200-probe signature, being more highly expressed in CHOP-resistant lymphomas (10.1 vs. 9.1 median RMA expression value, p = 0.0001, Wilcoxon rank sum test). This is notable because MYC+ DLBCLs have inferior outcomes to standard chemotherapy compared to other DLBCLs (168). Single genes which are known to modulate chemotherapy sensitivity in the *Eμ-Myc* model (which include *Top2a*, *Akt*, *Bcl-2*, *p53*, *Pten*, and *p19*^{ARF} (83, 146, 149-152) were not present in the CHOP response signature, but two probes for AKT3 and three probes for PTEN were differentially expressed between CHOP sensitive and CHOP resistant lymphomas at nominal p-values between 0.01 to 0.05 (Wilcoxon rank sum test). When this genomic signature is applied to an independent dataset of DLBCL patients treated with either CHOP or CHOP with rituximab (R-CHOP) (135), the signature significantly separated groups of patients with distinct responses to these two regimens (Figures 7B and 7C). We next evaluated the genomic signature of CHOP resistance in the *Eμ-Myc* lymphoma samples. As shown in Figure 7D, Cluster 2 lymphomas were predicted to be more chemotherapy sensitive than Cluster 1 lymphomas ($p < 0.0001$, Wilcoxon rank-sum test).

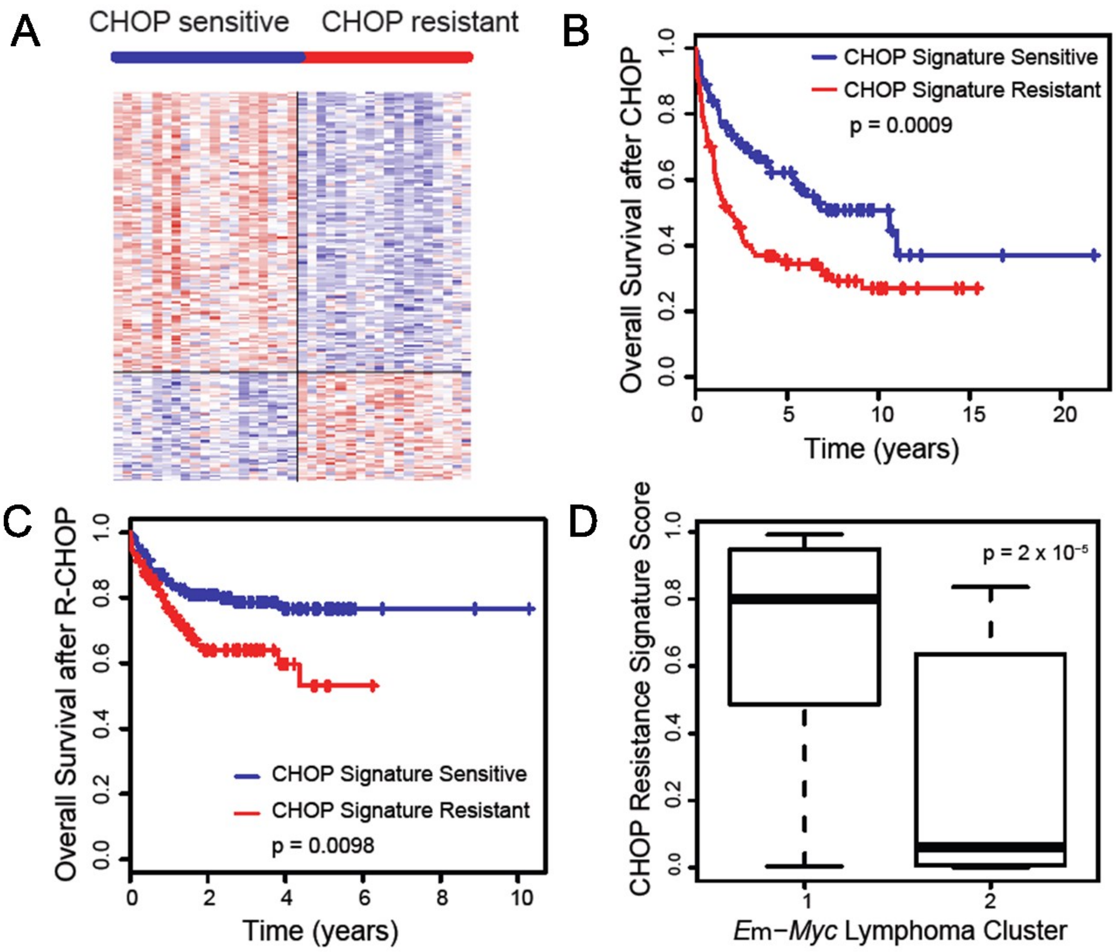


Figure 7. Differential predicted sensitivity to CHOP regimen

A) Heatmap of the genomic signature that discriminates human lymphomas with the longest remission after CHOP chemotherapy (“sensitive”) to those with the shortest remission (“resistant”). Red color represents upregulated gene probes and blue color represents downregulated gene probes. Overall survival of DLBCL patients after treatment with CHOP (B) or R-CHOP (C), grouped by CHOP signature score. P value calculated by log-rank test. D) Box and whisker plot of CHOP signature scores in Cluster 1 and Cluster 2 *Eμ-Myc* lymphomas. Parameters as described in Figure 3A. p-value calculated by Wilcoxon rank-sum test.

To validate the genomic predictions of differential sensitivity to chemotherapy between Cluster 1 and Cluster 2 *Eμ-Myc* lymphoma, we tested selected *Eμ-Myc* lymphoma clones for response to single agent doxorubicin and cyclophosphamide *in vivo*. These studies rely on the transplantability of *Eμ-Myc* lymphomas into immunocompetent C57BL/6 background strain mice. We selected lymphomas (C1-1, C1-3, C1-4, C1-5, C2-1, and C2-2) that were representative of the two clusters and that retained their characteristics after transplantation into recipient mice. These characteristics include genomic signature scores (Figure 8A), expression of B cell surface markers (Figure 8B), clonality as defined by single immunoglobulin heavy chain rearrangement (Figure 8C), and MYC, p53, p19^{ARF}, and MDM2 protein expression (Figures 8D and 8E). As such, the transplanted tumors appear to represent valid models of the two forms of *Eμ-Myc* lymphoma.

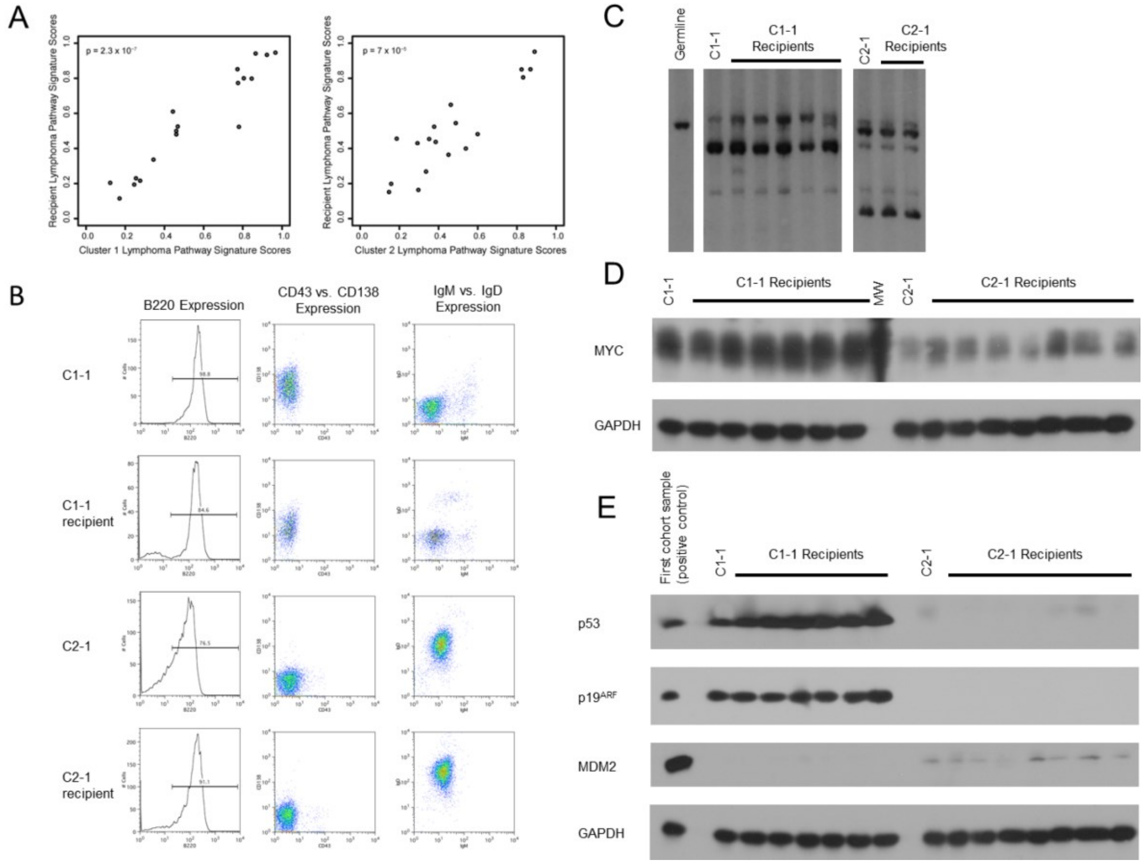


Figure 8. Consistency between original *Eμ-Myc* lymphoma and recipients

A) Cellular and oncogenic pathway activity signature scores from Cluster 1-1 lymphoma (left graph) or Cluster 2-1 lymphoma (right graph) plotted against scores from lymphomas derived in recipient mice from these lymphomas. P value calculated from Pearson's correlation. These findings indicate that transcriptional programs of pathway activity are not altered by transplantation into a new mouse. B) Surface expression of B220, CD43, CD138, IgM, and IgD in C1-1 and C2-1 lymphomas and their representative recipients, demonstrating stability in expression of these markers between the primary and the transplanted lymphomas. C) Southern blots probing for size of JH3-JH4 B-cell receptor fragment in germline DNA: C1-1 and its transplanted recipient lymphomas, and C2-1 and its transplanted recipient lymphomas. This result shows that these monoclonal lymphoma clones are maintained after transplantation. D) Immunoblot for MYC protein expression in C1-1 and C2-1 and the transplanted recipient lymphomas, demonstrating maintained MYC expression after transplantation. GAPDH expression demonstrates equal protein loading. E) Immunoblots for p53, $p19^{ARF}$, and MDM2 protein expression in C1-1 and C2-1 and the transplanted recipient lymphomas, demonstrating stability in p53 axis perturbations after transplantation. GAPDH expression demonstrates equal protein loading.

Eμ-Myc lymphoma recipients were treated with single agent doxorubicin or cyclophosphamide with one-time dosing, as previously described (149). As shown in Figure 9, therapeutic responses correlated with genomic signature predictions of CHOP resistance: mice bearing the Cluster 2-type *Eμ-Myc* lymphoma had significantly longer responses with either of the single agent conventional chemotherapy agents than mice bearing the Cluster 1-type lymphoma. After doxorubicin treatment, mice bearing the Cluster 1 lymphomas C1-1 (n = 7), C1-3 (n = 5), C1-4 (n = 5), or C1-5 (n = 6) continued to have ill appearance and rapid progression of lymphoma. On the other hand, almost all the mice bearing the Cluster 2 lymphomas C2-1 (n = 4) or C2-2 (n = 4) improved to normal appearance for approximately two weeks prior to progression. After cyclophosphamide therapy, mice bearing the Cluster 1 lymphomas C1-1 (n = 9), C1-3 (n = 7), C1-4 (n = 6), or C1-5 (n = 6) returned to a healthy appearance for approximately three weeks prior to progression. However, mice bearing the Cluster 2 lymphomas C2-1 (n = 5) or C2-2 (n = 4) returned to a fully healthy appearance for approximately five weeks prior to progression. As a negative control, mice bearing either Cluster 1 or Cluster 2 lymphomas that were not treated had progressive disease and were sacrificed one to two days from when treatment would have been administered. Upon dissection, there was a severe burden of lymphoma in all mice (data not shown).

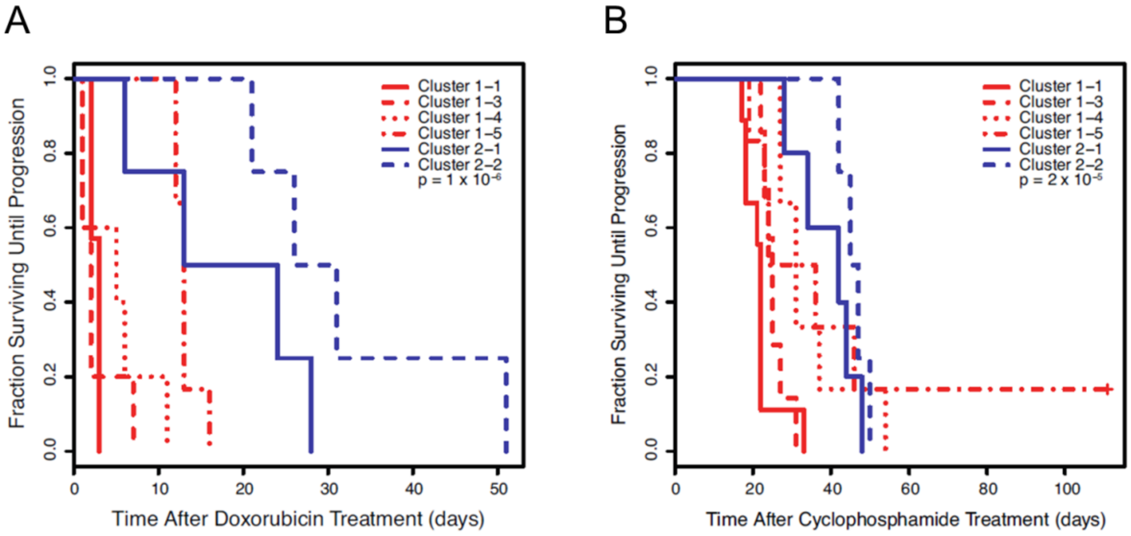


Figure 9. ***Eμ-Myc*** **lymphoma response to chemotherapy**

A) Progression free survival in recipient mice bearing Cluster 1 lymphomas (red) compared to Cluster 2 lymphomas (blue) after treatment with doxorubicin. B) Progression free survival in recipient mice bearing Cluster 1 lymphomas (red) compared to Cluster 2 lymphomas (blue) after treatment with cyclophosphamide. p-values calculated using the log-rank test.

The selected *Eμ-Myc* lymphoma clones are representative of the lymphoma clusters in terms of predictions of chemotherapy response, based on the CHOP resistance genomic signature. The Cluster 1-type lymphomas C1-1, C1-3, C1-4, and C1-5 had a CHOP resistance signature score of 0.8, 0.97, 0.94, and 0.98 respectively, while the Cluster 2-type lymphoma C2-1 and C2-2 had a CHOP resistance signature score of 0.35 and 0.02. These scores correlate with the chemotherapy responses shown in Figure 9.

Since perturbations of the p53 tumor suppressor axis are associated with inferior response to conventional chemotherapy in aggressive human lymphomas (169, 170) and in *Eμ-Myc* lymphomas (83), we also evaluated the response to chemotherapy based on the status of p53. As noted above, each of the Cluster 2 lymphomas exhibited wild type *Trp53* whereas the majority of the Cluster 1 lymphomas contained a mutant or deleted *Trp53* gene. Nevertheless, there were examples of Cluster 1 lymphomas that did contain a wild type *Trp53* gene, and as seen in Figures 10A and 10B, the response to chemotherapy for a Cluster 1 *Trp53* wild type was more similar to the Cluster 1 with mutant *Trp53* rather than the Cluster 2 lymphomas. This suggests that p53 status is likely not the sole determinant of response to chemotherapy in this model.

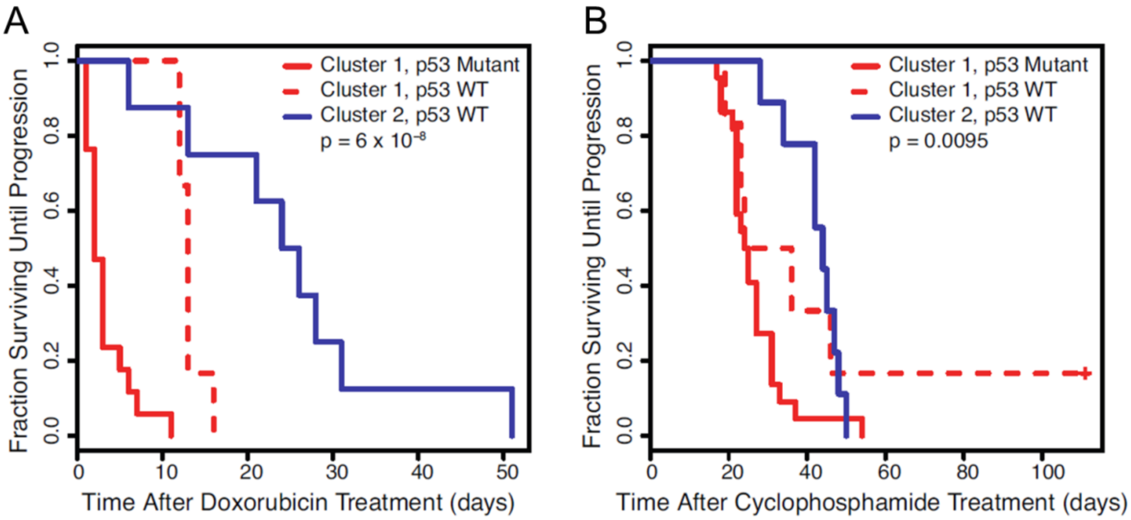


Figure 10. p53 status and chemotherapy response

Progression free survival in recipient mice bearing Cluster 1 lymphomas that have p53 mutation or deletion (red solid line) compared to Cluster 1 lymphomas that are p53 wild-type ("WT," red dashed line) and Cluster 2 lymphomas (blue solid line) after treatment with doxorubicin (A) and cyclophosphamide (B). p-values calculated using the log-rank test.

Together, these findings support the concept that variation in therapy response exists in transgenic *Eμ-Myc* mice. Additionally, transcriptional programs shared between human aggressive B-cell lymphomas and *Eμ-Myc* mice appear to serve as predictive biomarkers of response to conventional chemotherapy. As such, *Eμ-Myc* mice have utility as a model for variation in therapeutic response in human lymphomas, and for this reason, *Eμ-Myc* mice might be able to identify novel targets to rationally select and design treatment for human aggressive B-cell lymphomas.

2.3.4 Evaluating pathway specific targeted therapies in the context of *Eμ-Myc* lymphoma heterogeneity

Our prior work and the work of others defined gene expression signatures of cellular pathway activity and demonstrated that these signatures can serve as predictive biomarkers of response to pathway-specific targeted therapies (20, 21, 171). As shown in Figure 11A, we found significant distinctions in cellular pathway activity between types of human aggressive B-cell lymphomas, with RAS, MYC, PI3K, and E2F1 pathways significantly upregulated in BL, while TGFβ, STAT3, TNFα, EGFR, and interferon pathways are significantly upregulated in DLBCL (non-molecular BL). Visually, a similar pattern of pathway activity is seen between Cluster 1 and Cluster 2 *Eμ-Myc* lymphoma samples as is seen between BL and DLBCL (Figure 11B). To quantitate the similarity, we calculated binary logistic regression coefficients of the genomic signatures with respect to the human and *Eμ-Myc* lymphoma, and found a significant correlation between the coefficients for the lymphomas (cor = 0.961, Pearson's correlation test, Figure 11C).

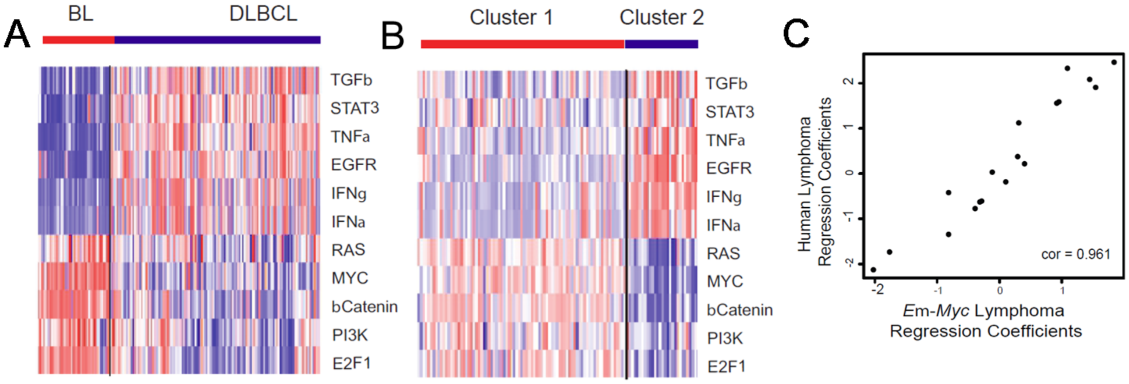


Figure 11. Patterns of pathway activation in *Eμ-Myc* lymphomas and human disease

A) Heatmap of scores from cellular and oncogenic pathway activity signatures in molecularly defined Burkitt lymphoma (BL) compared to diffuse large cell lymphoma (DLBCL). Red represents upregulated pathways and blue represents downregulated pathways. B) Heatmap, with color scheme as in Figure 11A, comparing Cluster 1 and Cluster 2 *Eμ-Myc* lymphomas from the training set. C) Binary regression coefficients from fitting each pathway activity signature scores to the phenotype of BL vs. DLBCL are plotted against binary regression coefficients from fitting each pathway activity signature score to the phenotype of Cluster 1 vs. Cluster 2, and were compared using Pearson's correlation.

Our goal in this research is to use these genomic signatures of pathway activity to identify novel therapeutic targets for cancers. To initiate these studies in lymphoma, we began by evaluating the NF-κB pathway, a known oncogenic pathway in DLBCL. We evaluated a genomic signature of the NF-κB pathway in human and *Eμ-Myc* lymphoma, and thereafter tested the extent to which therapeutic inhibition of the NF-κB pathway is beneficial in specific subgroups of lymphoma.

In human aggressive B-cell lymphomas, the NF-κB pathway appears to have divergent functions. For example, in DLBCL, NF-κB pathway is a well-recognized pro-survival and oncogenic mechanism (172), while in BL, the NF-κB pathway appears to act in a pro-apoptotic fashion within the context of *MYC* overexpression (173). The central role of the NF-κB pathway in DLBCL has led to the clinical evaluation of bortezomib, a proteosome inhibitor that suppresses the NF-κB pathway, where there appears to be specific efficacy in the ABC subtype of DLBCL (174, 175). Because of the variable role and activity of the NF-κB pathway in the spectrum of human aggressive B-cell lymphomas, we hypothesized that the same might be seen in *Eμ-Myc* lymphomas.

To assess NF-κB pathway activity in *Eμ-Myc* lymphomas, we utilized a genomic signature of the "TNFα pathway," which was developed from gene expression data obtained from endothelial cells treated with TNFα (GSE9055(176)) and is known to reflect activation of the NF-κB pathway. Using this genomic signature, we found significantly higher predictions of NF-κB pathway activity in the Cluster 2 *Eμ-Myc* lymphomas compared to the Cluster 1 lymphomas (Figure 12A, $p < 0.0001$). The predictions of differential NF-κB pathway activity between the two lymphoma clusters is supported by finding elevated protein expression of phosphorylated p65/RELA in Cluster 2 lymphomas compared to Cluster 1 (Figure 12B and Table 7) and by differential mRNA expression of NF-κB target genes between the two lymphoma clusters (Table 8).

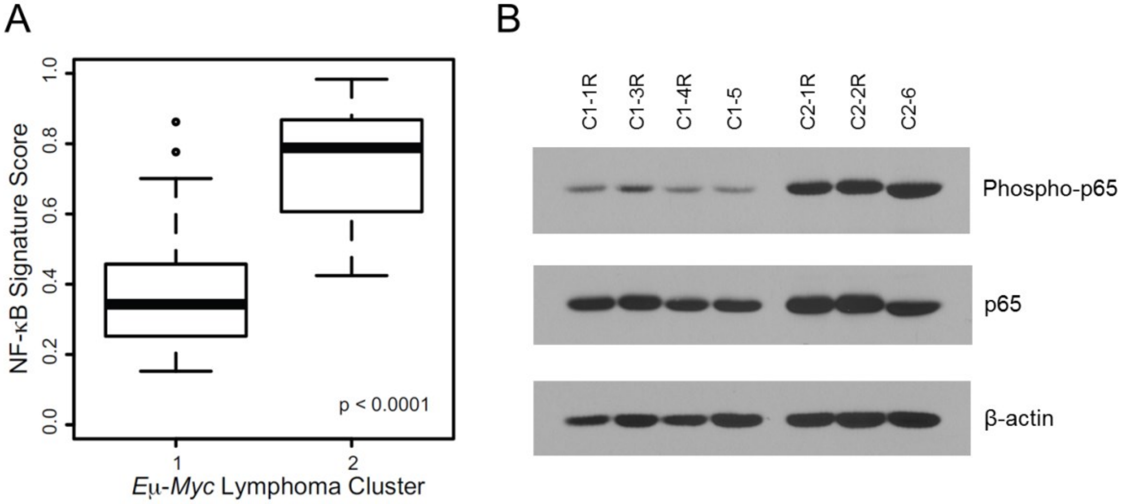


Figure 12. NF-κB pathway in *Eμ-Myc* lymphomas

A) Box and whisker plot of the NF-κB signature scores comparing Cluster 1 and Cluster 2 *Eμ-Myc* lymphomas. Bold line represents median, box represents the interquartile range (IQR), whiskers represent 1.5 times the IQR, and dots represent outliers. p-value calculated using the Wilcoxon rank sum test. B) Immunoblots for phosphorylated p65 (RELA) and total p65 protein expression in six *Eμ-Myc* lymphomas (three Cluster 1 and three Cluster 2) showing a higher level of phosphorylated p65 in Cluster 2 *Eμ-Myc* lymphomas, supporting the prediction of higher NF-κB activity in this group of lymphomas.

Table 7: *Eμ-Myc* lymphoma TNFa signature

Lymphoma	TNFa signature score
C1-1	0.35
C1-3	0.25
C1-4	0.18
C1-5	0.29
C2-1	0.84
C2-2	0.80
C2-6	0.90

Table 8: mRNA expression of selected NF-κB target genes in Cluster 1 and Cluster 2 *Eμ-Myc* lymphomas

Gene Name	Affymetrix Probe ID	Median Cluster 1 expression	Median Cluster 2 expression	P-value (Wilcoxon rank-sum test)
Tnfsf13b (Baff)	1460255_at	5.7027	6.5399	1.976321e-06
Blimp1 (Prdm1)	1420425_at	5.5224	7.1433	1.876977e-11
Ccl4	1421578_at	6.6616	7.336	2.346668e-10
Cd40	1439221_s_at	6.2323	6.381	0.005841444
Il2ra	1420692_at	4.8648	5.2171	2.540128e-05
B2m	1449289_a_at	12.9391	13.464	6.416793e-08
Cd44	1452483_a_at	7.9538	8.6154	1.975549e-05
Abca1	1421839_at	6.2129	6.7607	1.581027e-08
Fas	1460251_at	5.5219	5.8393	0.0002021305
Pim1	1423006_at	7.1451	7.9762	4.425071e-11

Because of the differential NF-κB pathway activity in the two *Eμ-Myc* lymphoma clusters, we hypothesized that therapeutic inhibition of the NF-κB pathway would be more beneficial in mice bearing Cluster 2 *Eμ-Myc* lymphomas than in mice bearing Cluster 1 *Eμ-Myc* lymphomas. As bortezomib is known to inhibit the NF-κB pathway (177) and has been evaluated in human aggressive lymphomas, we tested *Eμ-Myc* lymphoma Cluster 1 and Cluster 2 clones for their response to bortezomib.

As seen in Figure 13A, single agent bortezomib therapy was significantly more effective in mice bearing the Cluster 2 *Eμ-Myc* lymphomas than Cluster 1 *Eμ-Myc* lymphomas. The recipients of Cluster 2 lymphomas C2-1 (n = 6) and C2-2 (n = 8) treated with bortezomib had stabilization of disease for approximately one to two weeks, while recipients of Cluster 1 lymphomas C1-1 (n = 6), C1-3 (n = 6), C1-4 (n = 6), and C1-5 (n = 6) treated with bortezomib almost uniformly progressed within days of the first injection. As with cyclophosphamide and doxorubicin, Cluster 2 lymphomas (all *Trp53* wild type) had longer progression free survival than Cluster 1 lymphomas (either *Trp53* wild type or mutant/deleted) after bortezomib treatment (Figure 13B). The NF-κB pathway signature scores for the Cluster 2 lymphomas (0.84 and 0.8 respectively) were higher than the NF-κB pathway signature scores for the Cluster 1 lymphomas (0.35, 0.25, 0.18, and 0.29 respectively), and the scores correlated with the clinical responses seen.

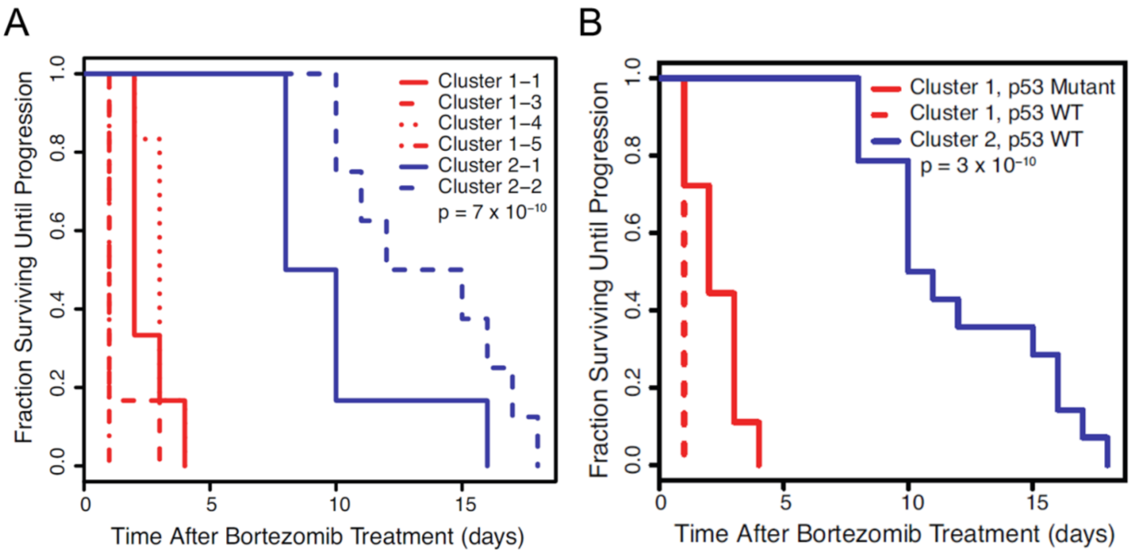


Figure 13. ***Eμ-Myc*** **response to bortezomib**

A) Progression free survival in recipient mice bearing Cluster 1 lymphomas (red) compared to Cluster 2 lymphomas (blue) after treatment with bortezomib. p-values calculated using the log-rank test. B) Progression free survival in recipient mice bearing Cluster 1 lymphomas that have p53 mutation or deletion (red solid line) compared to Cluster 1 lymphomas that are p53 wild-type ("WT," red dashed line) and Cluster 2 lymphomas (blue solid line) after treatment with bortezmib.

The results of applying and testing genomic signatures of pathway activity to the *Eμ-Myc* model demonstrate that they can be used as biomarkers of therapeutic response. Since *Eμ-Myc* lymphoma is a model for heterogeneity seen in human aggressive B-cell lymphomas, and since mouse and human lymphomas have similar patterns of pathway activity as measured by genomic signatures, we believe that *Eμ-Myc* mice can serve as a model to identify and test novel targeted therapies for human aggressive B-cell lymphoma subtypes.

2.4 Discussion

Perhaps the greatest challenge facing the development of new drugs and effective treatment for human cancers is the heterogeneity of the disease. Cancer in general is not one disease, and the same holds true for B-cell lymphomas. BL, DLBCL, and B-cell lymphomas with features intermediate between DLBCL and BL represent a spectrum that exhibits clinical heterogeneity and diversity of underlying genetic alterations. An ability to define and understand this heterogeneity is clearly critical to adapting and applying effective treatment strategies to individual patients.

Equally important to attaining the goal of personalized lymphoma therapy is the development of disease models that can be used for the evaluation of new therapeutic strategies. The *Eμ-Myc* transgenic mouse was developed as an example of B-cell lymphoma initiated by the *Myc* oncogene, and various studies have used this model to investigate how the manipulation of cooperating genes can alter time to lymphoma onset or modulate response to chemotherapy (74, 83, 144, 146, 149-151). However, the naturally inherent heterogeneity of the model is not well appreciated.

Our prior studies highlighted the fact that the *Eμ-Myc* model develops a variety of genomically diverse lymphomas that occur naturally and have distinct differences in time to lymphoma onset (143). Here, we confirm these findings and develop methods that can be used to classify new lymphomas and link them to distinct human lymphoma subtypes. A key and critical component of our approach is the capacity to use gene expression profiling as a common currency to allow for cross-species comparisons (69). Genomic signatures developed in the human context and applied to mouse data, and vice versa, demonstrate similarities in biology between human and murine lymphomas. Even though all *Eμ-Myc* lymphomas derive from *Myc*-overexpressing B-cells and human B-cell lymphomas do not necessarily have *MYC* aberrations, our approach shows that additional genomic programs connect human and murine lymphomas.

Although the *Eμ-Myc* model shares genomic programs with aggressive human B-cell lymphomas, it does have certain limitations. Recent next generation sequencing efforts have identified numerous mutations that are found with varying frequency in human aggressive lymphomas (140-142). Although we do not know the extent to which the *Eμ-Myc* model recapitulates the full breadth of genetic alterations in human disease, the resulting gene expression patterns demonstrate the downstream similarities between human and murine lymphomas. In addition, we note differences comparing the subtypes of *Eμ-Myc* and human aggressive lymphoma, such as presentation (localized vs. extensive disease), cell of origin determination (GCB vs. ABC subtypes), and response to chemotherapy. These differences highlight that the mouse model is not an exact representation of aggressive human lymphomas. Despite these limitations, we believe our approach provides the rationale for using *Eμ-Myc* mice and genomic signatures to model heterogeneity of therapy response in aggressive human B-cell lymphomas.

Most importantly, we now shift the focus of these studies from purely descriptive to one where *Eμ-Myc* mice can serve as a model to test therapies that may have future utility in human aggressive lymphomas. Our evaluation of conventional chemotherapy drugs used in the treatment of lymphoma validates that *Eμ-Myc* mice can serve to model heterogeneity in response to therapy in the human setting. Aberration in the p53 gene, resulting in abnormal protein expression, is known to modulate the response of lymphomas to conventional chemotherapy. Our evaluation of the p53 axis in *Eμ-Myc* mice confirms these findings since all of the p53 mutations were found in Cluster 1 lymphomas, which displayed the greatest resistance to chemotherapy. Nevertheless, our work suggests that additional genetic programs may affect chemotherapy response in the *Eμ-Myc* mouse model since one cluster 1 lymphoma with wild type p53 also exhibited chemotherapy resistance. However, this limited analysis with only one such clone precludes conclusions regarding the relative roles of p53 status versus expression-based clusters in predicting therapeutic response. Further experiments with additional lymphoma clones as well as with other conventional chemotherapy agents will be necessary to fully assess this.

Our previous work and that of others have established a link between pathway activation and sensitivity to pathway-specific therapeutic agents (20, 171). Using the *Eμ-Myc* model to evaluate targeted therapeutics in different subtypes of lymphoma holds great promise. In experiments beginning to explore these possibilities, we observed that treatment with bortezomib, a proteosome inhibitor that suppresses NF-κB pathway activity, had preferential benefit in treating lymphomas derived from the Cluster 2 subgroup, which we demonstrate has having similarities to the ABC-subtype of DLBCL. These data confirm results emerging from human clinical trials regarding the use of bortezomib-containing regimens in the ABC subgroup of DLBCL (174, 175). Moving forward, our results suggest that predictions of pathway activity could be used to select and test new therapeutic options and combinations in an *in vivo* experimental model that reflects characteristics of subtypes of human lymphoma. Ultimately, such findings could inform patient stratification and design of future clinical trials.

3. Genomic analysis of the role of NFYB in E2F1-mediated apoptosis

3.1 Introduction

E2Fs are a family of transcription factors important for the regulation of cell proliferation and apoptosis. E2F activation, resulting from cyclin-dependent kinase inhibition of retinoblastoma (Rb) protein function, is the trigger that leads to the transition from G_0 to G_1-S and initiation of the cell cycle. Genetic lesions in the Rb tumor suppressor pathway lead to unrestrained E2F activity and deregulated cell proliferation, critical to the development of numerous cancers (178-180). Among the three activator E2Fs, E2F1-E2F3, E2F1 is unique in its ability to induce apoptosis as well as proliferation (181). Overexpression of E2F1 in quiescent fibroblasts induces apoptosis (182) while E2F1-/- mouse thymocytes are resistant to apoptotic stimuli (183). The induction of numerous apoptotic genes and the repression of survival genes have been documented during E2F1-dependent apoptosis (120, 123, 184-187). Additionally, DNA damage leads to E2F1 stabilization through ATM -mediated phosphorylation and activation through PCAF-mediated acetylation, resulting in apoptosis (123, 124). Together, these studies suggest an important role of E2F1 deregulation in cancer as well as the therapeutic potential to harness the apoptotic activity of E2F1 for cancer therapy. However, better understanding of the control mechanisms balancing the proliferative and apoptotic activities of E2F1 is necessary to realize this potential.

Previous studies have demonstrated that E2F1 induces numerous other transcription factors, some of which cooperate with E2F1 during induction of target genes. For example, E2F1 induces and then cooperates with FOXO1/3 to induce apoptotic genes (128). In contrast, E2F1 induces TopBP1, which binds to E2F1 and specifically represses its apoptotic activity (188). These examples underscore the importance of other transcriptional regulators in determining the transcriptional and phenotypic outcomes of E2F1.

The nuclear transcription factor Y (NF-Y) recognizes and binds to the CCAAT box, which is highly enriched in the promoters of genes regulated during the G_2/M phase (189, 190). NF-Y is a trimeric complex that is composed of NFYA, NFYB, and NFYC subunits, all of which are necessary for DNA binding. NFYB and NFYC contain a histone-like motif and form a dimer which is required for association with NFYA and sequence-specific DNA binding (191). The NF-Y complex plays a role in regulating proliferation by controlling expression of genes required for cell cycle progression such as cyclin A, cyclin B2, CDC25A, CDC25C, and CDK1 (192-195). Furthermore, NF-Y regulates cell survival through direct control of several anti-apoptotic genes (196, 197). Paradoxically, ectopic expression of NFYA was recently shown to induce apoptosis through upregulation of E2F1 (198).

In this chapter, I will present data that support an inverse relationship, whereby E2F1 induces the expression of NFYB, which, together with E2F1, regulates a large group of joint target genes. These NFYB-dependent E2F1 target genes including many apoptotic genes that, based on gene expression, contribute to a pro-survival role for NFYB in E2F1-mediated apoptosis. Furthermore, we show that overexpression of these genes occurs in sarcomas compared to normal control tissues and associates with chemotherapy resistance. Taken together, our results identify NFYB as a new important pro-survival player in E2F1 transcriptional program.

3.2 Materials and Methods

3.2.1 Cell culture

U2OS human osteosarcoma cells stably expressing ER-HA-E2F1 were grown in Dulbecco's Modified Eagle Medium with 10% fetal bovine serum (FBS) and 2.5ug/ml puromycin. U2OS cells stably expressing ER-HA-E2F1 and overexpressing NFYB were grown in the same media along with 10ug/ml blasticidin. 4-Hydroxytamoxifen (OHT) was obtained from Sigma (T176) and dissolved in ethanol.

3.2.2 RNA isolation and RT-PCR

Total RNA was isolated using RNeasy Kit (Qiagen) on a QIAcube (Qiagen). RNA quality and concentration was assessed using Nanodrop ND-1000 Spectrophotometer. A total of 1ug RNA was reverse transcribed using High-Capacity cDNA Reverse Transcription Kit (Life Technologies 4368814) according to manufacturer's instructions. cDNA was diluted 1:4 with water. cDNA was analyzed in triplicate by real time PCR on

Life Technologies StepOnePlus Real-Time PCR Systems with Life Technologies Power SYBR Green Master Mix (4367659).

3.2.3 Chromatin immunoprecipitation

ER-HA-E2F1 chromatin immunoprecipitations were conducted using the EZ-ChIP Kit from Millipore (17-371) according to the manufacturer's instructions using monoclonal anti-HA antibody (Covance 16B12). NFYB (GPH1017391(-)01A), SFRP1 (GPH1026162 (+)01A), PRKCZ (GPH1000044(-)01A), FOLH1 (GPH1016343(-)01A), and IGX1A (GPH100001C(-)01A) EpiTect ChIP primers were obtained from Qiagen. Immunoprecipitated DNA was analyzed in quadruplicate by real time PCR on an ABI Prism 7900HT Sequence Detection System with RT^2 SYBR Green ROX qPCR Mastermix (Qiagen 330520).

3.2.4 Microarray analysis

RNA for microarray analysis of U2OS ER-HA-E2F1 cells was prepared using the RNeasy Kit (Qiagen) on a QIAcube (Qiagen). RNA quality and integrity was analyzed using Agilent Lab-on-a-Chip RNA Bioanalyzer. RNA was amplified using Ambion Message-Amp Premier Kit and analyzed on Affymetrix U133A 2.0 microarrays by the Duke Microarray Facility. Raw microarray data was normalized using the MAS5 algorithm or RMA algorithm. Expression data is available in the Gene Expression Omnibus database under accession number GSE61272.

Unsupervised clustering was performed using Cluster 3.0. The MAS5 normalized data was log transformed, mean-centered by gene, and normalized by gene. Hierarchical clustering was performed by clustering genes using correlation (uncentered) similarity metric and complete linkage. Clustering results were viewed and figures generated using Java Treeview.

Pathway signatures were generated using the ScoreSignature module in GenePattern[4] using MAS5 and RMA normalized microarray data. Pathway signatures were clustered by pathway in Cluster 3.0 using correlation (uncentered) similarity metric and complete linkage. Heatmap was generated using Matlab software.

3.2.5 siRNA and lentivirus infection

siRNA reverse transfections were performed using Lipofectamine RNAiMax reagent (Life Technologies 13778150). Negative control siRNAs were from Sigma (SIC001) and Dharmacon (D-001810-01). siRNAs targeting NFYB were from Dharmacon (J-010002-08) and Ambion (4392420).

V5-tagged NFYB driven by CMV promoter overexpression vector was obtained from DNASU (HsCD00329598). Lentiviral particles were produced using 293T cells co-transfected with the lentiviral vector, pMD2.G (Addgene 12259), and psPAX2 (Addgene 12260) using Mirus TransIT-LT1 Transfection Reagent (MIR 2300). U2OS ER-HA-E2F1 cells were infected followed by selection with blasticidin (10ug/ml) for two weeks.

---

[4] http://www.broadinstitute.org/cancer/software/genepattern/

3.2.6 Western blotting

Cell pellets were lysed with RIPA buffer containing Complete Mini protease inhibitor cocktail (Roche) and PhosSTOP phosphatase inhibitor cocktail (Roche). Proteins were resolved on 10% SDS-PAGE, transferred to Immobilon PVDF membranes and probed with antibodies against GAPDH (Santa Cruz, Sc-25778), NFYB (Santa Cruz, Sc-13045), cleaved PARP (Cell Signaling, 9541S), and V5 (Pierce, MA5-15253). Equal protein loading was verified by probing for GAPDH.

3.2.7 Statistical analysis

Graphs and significance levels were generated using Graphpad software. Pooled results are presented as mean and standard deviation of triplicate experiments for expression-based real time PCR and quadruplicate experiments for chromatin immunoprecipitation-derived real time PCR. For determining the list of genes regulated by NFYB, p values were calculated using two tailed t-test with unequal variance.

3.2.8 Oncomine analysis

The list of genes generated from the microarray analysis (Appendix A) was entered into Oncomine Research Premium Edition as a custom concept and analyzed for associations across all Oncomine datasets with the following thresholds: p-value of 1E-4 and odds ratio of 2.

3.3 Results

3.3.1 NFYB is a direct target of E2F1

Previous work has documented the role of E2F transcription factors in the control of gene expression central to cell cycle and cell fate decisions (115). Further work has provided evidence for a complex combinatorial mechanism of gene regulation involving E2F proteins together with other cooperating transcriptional regulators (127, 199, 200). This includes regulatory cascades in which E2Fs activate the expression of genes encoding other transcription factors and then cooperate with these induced factors to regulate additional target genes, a relationship know as a feed-forward regulatory loop (128, 201). To further explore this concept, with a focus on the E2F1 transcription factor, we have made use of an inducible system with an ER-E2F1 chimeric protein expressed in U2OS cells. Upon the addition of 4-hydroxytamoxifen (OHT), the chimeric protein translocates to the nucleus and activates E2F1-mediated transcription. These cells have been previously characterized showing minimal effect of OHT on the original parental cell line (106, 128).

In a recent study of E2F1 target genes, we found that genes encoding transcription factors represent over 13% of E2F1 target genes with “Regulation of Transcription” as a significantly enriched Gene Ontology (GO) category ($p<0.0001$) (128). Among the transcription factors induced by E2F1, we have focused on NFYB since NFY proteins, similar to E2F1, are important regulators of both proliferation and apoptosis (reviewed in (202)).

To validate the microarray results showing induction of NFYB by E2F1, we analyzed NFYB expression levels by real-time PCR. Following OHT induction, NFYB expression increased linearly through four and eight hours to reach a level almost seven-fold higher than prior to E2F1 induction (Figure 14A). This induction of NFYB RNA is also reflected at the protein level by Western blotting where the induction of NFYB protein can be seen at eight and twenty four hours following OHT induction (Figure 14B).

To address whether NFYB induction is a direct effect of the action of E2F1, we performed a chromatin-immunoprecipitation (ChIP) analysis to determine whether E2F1 binds to the NFYB promoter. Following induction by OHT, E2F1 binding to the NFYB promoter increased more than four-fold while binding to IGX1, a negative control, did not change (Figure 14C). This suggests that activation of NFYB transcription results from direct binding of E2F1 to the NFYB promoter.

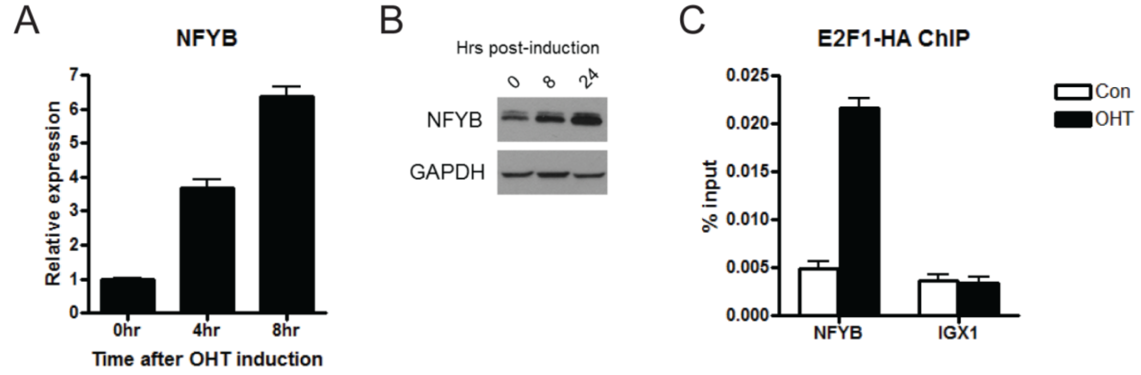


Figure 14. NFYB is a direct E2F1 target

A) Real-time PCR analysis of NFYB mRNA levels at four and eight hours after E2F1 induction in U2OS ER-E2F1 cells. Cells were serum starved for twenty four hours prior to addition of 80 nM OHT. B) Western blot of NFYB and GAPDH protein levels. U2OS ER-E2F1 cells were induced with 80nM OHT for eight and twenty four hours following twenty four hour serum starvation. Lysates were analyzed by SDS-PAGE/Western blot and probed with anti-NFYB and anti-GAPDH antibodies (loading control). C) U2OS ER-E2F1 cells were serum starved for twenty four hours followed by induction with 80nM OHT for seven hours. Chromatin immunoprecipitation was performed with anti-HA antibody for detection of ER-HA E2F1 binding to the NFYB promoter and IGX1 repeats (negative control).

3.3.2 Role of NFYB in E2F1-mediated transcriptional activation

To explore the role of NFYB in E2F1-mediated transcriptional activation, we knocked down the expression of NFYB using two different small interfering RNA (siRNAs) targeting NFYB. Efficient knockdown of NFYB mRNA levels was validated by real-time PCR (Figure 15A) and Western blot analysis (Figure 15B).

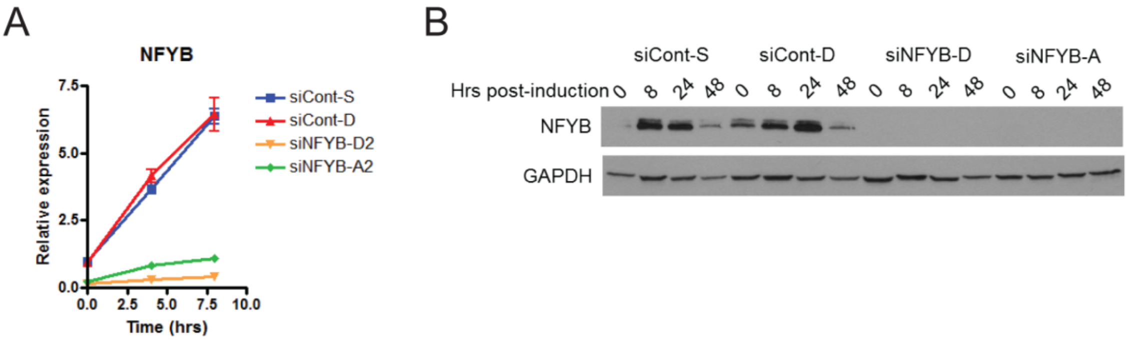


Figure 15. NFYB mRNA and protein knockdown

A) Real-time PCR analysis of NFYB mRNA levels following siRNA transfection. U2OS ER-E2F1 cells were transfected with two siRNAs targeting NFYB or two negative control siRNAs at 100nM. Cells were serum starved for twenty four hours followed by 80nM OHT induction for four or eight hours. B) Western blot analysis of NFYB protein levels following E2F1 activation. U2OS ER-E2F1 cells transfected with siRNA targeting NFYB or control siRNAs were serum starved for twenty four hours followed by OHT induction for eight, twenty four, and forty eight hours. Lysates were analyzed by SDS-PAGE/Western blot and probed with anti-NFYB and anti-GAPDH antibodies (loading control).

We next interrogated the global gene expression effects of NFYB knockdown on E2F1-mediated transcription using expression microarrays. Comparison of genome-wide expression levels after E2F1 activation in cells transfected with either negative control siRNAs or siRNAs targeting NFYB identified 197 genes whose expression in E2F1 activated cells (after 8 hours of OHT induction) was significantly different ($p<0.001$) following NFYB knockdown by at least 1.3 fold higher or 0.7 fold lower compared to control siRNA. As expression of these genes is affected by NFYB knockdown and given that E2F1 directly targets NFYB, these genes represent the broadest spectrum of genes directly or indirectly targeted by E2F1 and NFYB.

Unsupervised clustering of these genes reveals five clusters of genes: 1) induced by E2F1 activation and reduced by NFYB knockdown, 2) unaffected by E2F1 activation and reduced by NFYB knockdown, 3) reduced by E2F1 activation and further reduced by NFYB knockdown, 4) induced by E2F1 activation and further induced by NFYB knockdown, 5) reduced by E2F1 but higher in NFYB knockdown (Figure 16A and Appendix A).

Further examination of the genes within the clusters also revealed multiple genes involved in apoptosis and survival. For example, the genes within cluster 2, whose expression reduced by NFYB knockdown, include API5, an inhibitor of E2F1-mediated apoptosis, and MALT1, a paracaspase that promotes activation of NF-κB signaling (129, 203). Cluster 4, in which expression was induced by E2F1 and further increased by NFYB knockdown, includes SIVA1, apoptosis inducing factor, a proapoptotic gene that plays an important role in CD27-mediated apoptosis (204) and is known to be induced by E2F1 (205). We have used RT-PCR to validate the expression of API5 (Figure 16B) and SIVA1 (Figure 16C).

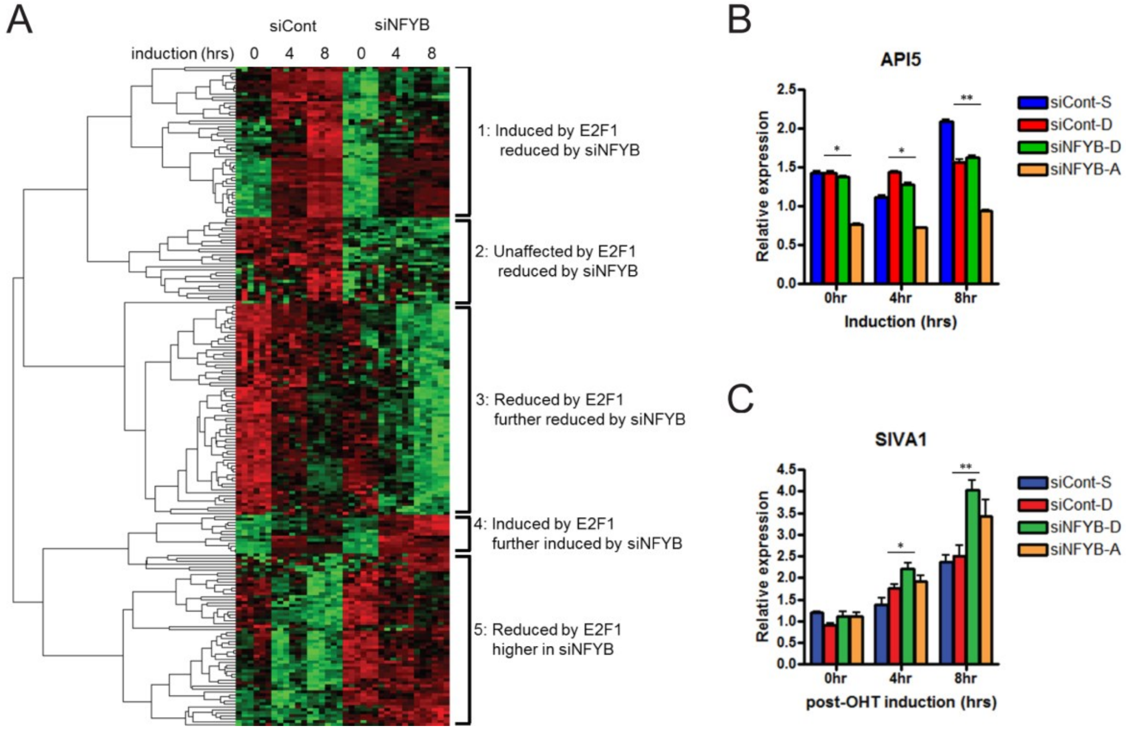


Figure 16. NFYB target genes and validation

A) Microarray analysis of the effect of NFYB knockdown on E2F1-mediated transcription. Samples were processed in the same manner as in Figure 2A and analyzed using Human U133A 2.0 expression microarrays. Heatmap represents the results of hierarchical clustering of 224 probes that showed at least 1.3-fold increase or 0.7-fold decrease in expression (significance of $p<0.001$) compared to control following NFYB knockdown at eight hours. B) Real-time PCR validation of target gene expression decrease for API5. C) Real-time PCR validation of target gene increase for SIVA1. Samples were processed in the same manner as Figure 15A. * denotes $p<0.05$, ** denotes $p<0.01$, *** denotes $p<0.001$, **** denotes $p<0.0001$.

In order to better understand the transcriptional program directly regulated by E2F1 that is dependent upon NFYB, we expanded our analysis of cluster one. Further analysis of microarray data identified 148 genes whose expression is induced at least two fold by E2F1 activation and is twenty percent higher (significance at $p<0.05$) in induced control siRNA samples compared to induced siNFYB samples (Figure 17A and Appendix B). To further validate these genes that may be co-regulated by E2F1 and NFYB, we used real-time RT-PCR to measure the expression of those genes in which the effect of NFYB knockdown on expression levels was most significant. Validation for selected genes is presented for SFRP1 (Figure 17B), FOLH1 (Figure 17C), and PRKCZ (Figure 17D).

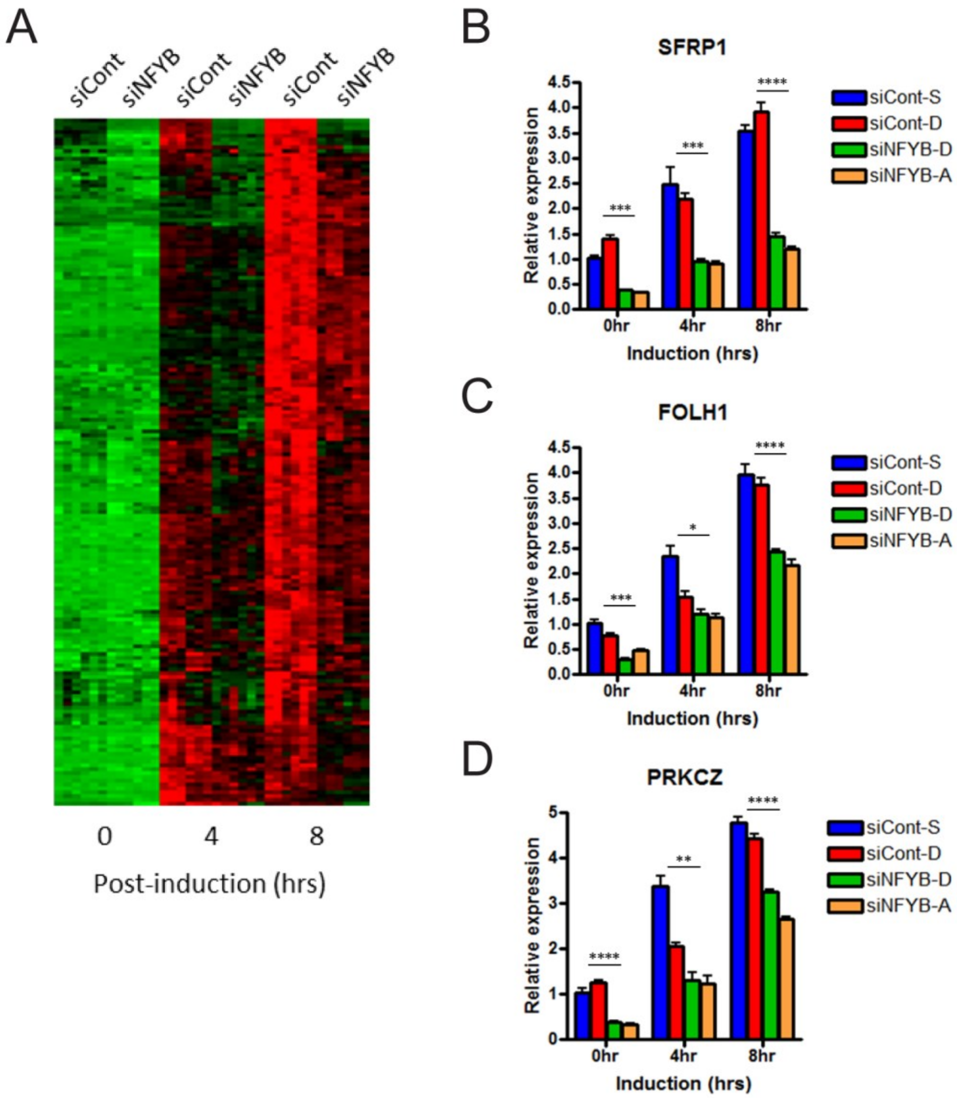


Figure 17. E2F1 and NFYB co-regulated genes

A) Heatmap represents expression levels of 174 probes that were upregulated at least two-fold after eight hours of E2F1 induction in control samples and levels of which were at least 20% lower (significant at $p<0.05$) in siNFYB than in si-control samples in eight hours samples. B) Real-time PCR validation of target gene expression decrease for SFRP1 following NFYB knockdown. C) Real-time PCR validation of target gene expression decrease for FOLH1 following NFYB knockdown. D) Real-time PCR validation of target gene expression decrease for PRKCZ following NFYB knockdown.

3.3.3 Effect of E2F1 and NFYB on predicted pathway activity

To determine what pathways are affected by E2F1 induction and NFYB knockdown, we performed a pathway analysis using the microarray data. Taking advantage of oncogenic, tumor suppressor, and tumor microenvironment signatures generated by the Nevins and Chi labs (20, 27), we assessed the likelihood that various pathways are activated (Figure 18A) using the ScoreSignature module in GenePattern.

The effect of E2F1 induction by OHT was validated with the E2F1 pathway predictions, showing low E2F1 activity prior to induction for both control and NFYB knockdown samples. In both groups, pathway activity increased dramatically following OHT induction, with peak induction seen at eight hours. Other pathways that paralleled this induction patterns were β-catenin (BCAT), p63, and SRC. Interestingly, BCAT and p63 show significantly higher predicted levels of activity in NFYB knockdown samples at zero and four hours for BCAT (Figure 18B) and all time points for p63. This suggests that NFYB has a strong effect on these pathways during the course of E2F1 induction. Other pathways showed an inverse effect, in which predicted pathway activity decreased following E2F1 induction. These pathways included EGFR, HER2, Myc, ER, PR, interferon α, and interferon β. However, NFYB knockdown has little effect on pathway status. Taken together, this data indicates that NFYB plays a role in pathway activity following induction by E2F1.

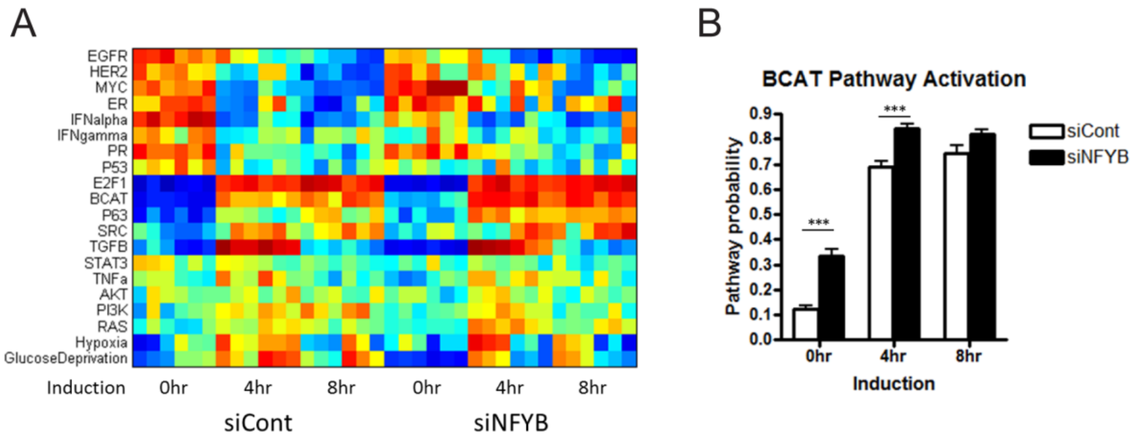


Figure 18. Patterns of pathway activation following NFYB knockdown

A) Heatmap of pathway signatures generated by ScoreSignature for all microarray samples. Blue denotes low probability and red denotes high probability. B) Predicted β-catenin pathway activity as compared between control and NFYB knockdown samples.

Earlier we showed that SFRP1 expression was induced following E2F1 activation and that this expression decreases following knockdown of NFYB. SFRP1 is a secreted protein that acts as an antagonist of the Wnt receptor (206), which under normal conditions activates the β-catenin pathway. Following NFYB knockdown, SFRP1 expression decreases and β-catenin pathway activity increases, suggesting the former may cause the latter. This is reflected in a comparison between predicted β-catenin pathway activity in control and NFYB knockdown samples, in which predicted activity is significantly higher in the latter group at zero and four hours and almost significantly higher at eight hours (p=0.06) (Figure 18B). This data suggests the NF-Y complex is involved in regulation of the β-catenin pathway through SFRP1. Interestingly, many other Wnt/β-catenin pathway genes were also present within the E2F1/NF-Y co-regulated gene set, hinting at the downstream effects of enhanced pathway activity.

3.3.4 NFYB-dependent E2F1 target genes

To establish whether the genes induced by E2F1 in an NFYB dependent manner are direct targets of E2F1, we performed chromatin immunoprecipitation (ChIP) analysis before and after induction by OHT. All three genes showed increased E2F1 binding to their promoters following E2F1 induction (Figure 19A). E2F1 recruitment specificity was demonstrated by the lack of E2F1 binding to the IGX1 negative control repeat sequences. These results indicate that E2F1 is physically associated with the regulatory regions of these genes to regulate their expression. As a complementary approach to further validate the involvement of NFYB in induction of these targets, we employed a gain of function approach. To this end, we infected the inducible U2OS ER-E2F1 cell line with a lentivirus containing NFYB cDNA. Following selection to establish a stable line, we validated NFYB ectopic expression by RT-PCR (Figure 19B) and Western blotting (Figure 19C).

To test whether increased expression of NFYB caused induction of the target genes, we induced the transgenic and parental line with OHT and compared target gene expression. Real-time PCR analysis demonstrated that overexpression of NFYB alone induced SFRP1 (Figure 19D) and FOLH1 (Figure 19E). Expression of these genes was also higher in NFYB-overexpressing cells following E2F1 induction. Taken together with the siRNA data, these results demonstrate that SFRP1 and FOLH1 are novel targets that are co-regulated by E2F1 and NFYB.

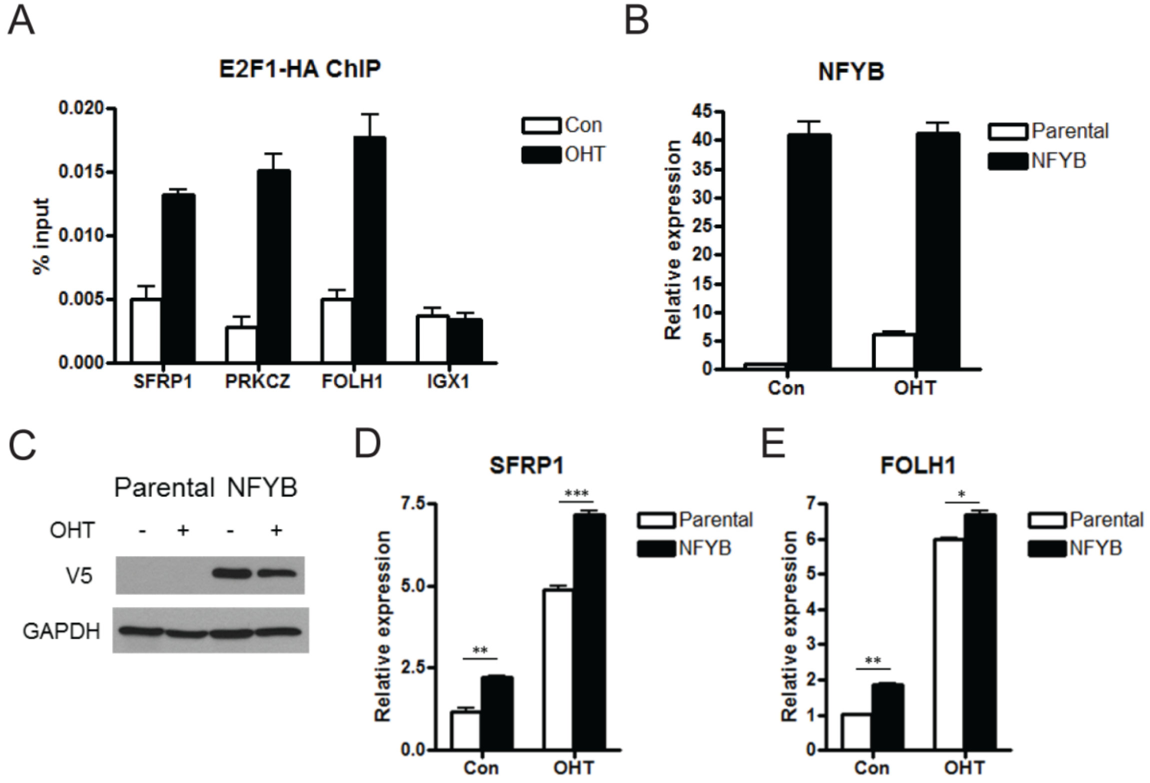


Figure 19. NFYB-dependent E2F1 targets

A) ChIP analysis of E2F1 binding to target gene promoters. U2OS ER-HA E2F1 cells were serum starved for twenty four hours followed by 80nM OHT induction for seven hours. Immunoprecipitation was performed with anti-HA antibody for detection of ER-HA E2F1 binding to the SFRP1, PRKCZ, FOLH1 promoters and to IGX1 repeats (negative control). B) Real-time PCR analysis of NFYB expression. U2OS ER-E2F1 cells were infected with lentviruses encoding the NFYB cDNA tagged with V5. A stable line was selected using blasticidin resistance for two weeks. This NFYB overexpression line is denoted as NFYB. The parental U2OS ER-E2F1 line is denoted as "parental". Cell lines were serum starved for twenty four hours followed by 80nM OHT induction for eight hours. C) Western blot analysis of ectopic NFYB expression following establishment of stable NFYB overexpression cell line. Lysates were analyzed by western blotting of NFYB transgene levels using an anti-V5 antibody, and GAPDH. Samples were processed in the same manner as in Figure 19B. D and E) Real-time PCR analysis of SFRP1 (D) and FOLH1 (E) in U2OS ER-E2F1 in parental and NFYB overexpressing U2OS ER-E2F1 cell lines. Samples were processed in the same manner as in Figure 19B.

Given the role that NFYB appears to play in regulating multiple anti-apoptotic genes, we sought to determine whether NFYB induction played a role in E2F1-mediated apoptosis. Western blot analysis following knockdown of NFYB by two different siRNAs demonstrated significantly increased E2F1-induced apoptosis as evidenced by the intensity of cleaved PARP protein band (Figure 20 and Fig 15B demonstrating knockdown efficiency). Coupled with our microarray data that identified multiple pro-survival NFYB target genes, this functional result suggests that induction of NFYB by E2F1 contributes to the attenuation of the apoptotic arm of E2F1 signaling.

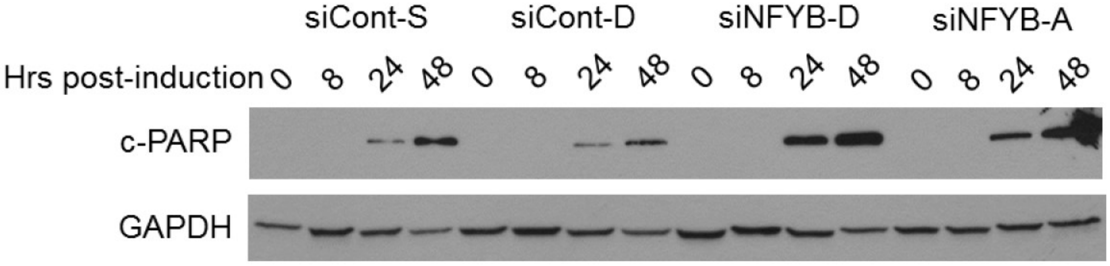


Figure 20. NFYB and E2F1-mediated apoptosis

Western blot analysis of cleaved PARP levels following E2F1 activation. U2OS ER-E2F1 cells transfected with siRNA targeting NFYB or control siRNAs were serum starved for twenty four hours followed by induction at eight, twenty four, and forty eight hours following OHT induction. Lysates were analyzed by SDS-PAGE/Western blot and probed with anti-cleaved PARP and anti-GAPDH antibodies (loading control). See Fig 15B for NFYB knockdown efficiency in these samples.

3.3.5 Genes from E2F1-NFYB signature are upregulated in sarcoma and drug resistant cell lines

In light of the established roles of E2F and NFY in cancer phenotypes, we sought to determine whether the transcriptional program jointly controlled by E2F1 and NFYB might define unique cancer states. Using the results from the microarray analysis, we generated an E2F1/NFYB signature comprised of 174 expression probes (148 genes) that were at least two fold induced by E2F1 and showed the largest decrease in expression following NFYB knockdown (Figure 17A and Appendix B). We made use of the Oncomine analysis suite, which integrates genome-wide expression and copy number data from more than 700 cancer-related datasets as a rich source of expression data coupled with phenotypic information (207). In five independent sarcoma datasets, the E2F1/NFYB signature is significantly associated with signatures (concepts) consisting of genes that are overexpressed in cancer compared to normal tissue. A representative heatmap shows the overexpression of genes from the E2F1/NFY signature in dedifferentiated liposarcoma compared to normal adipose tissue (Figure 21A). A second representative heatmap shows the overexpression of E2F1/NFY signature genes in gastrointestinal stromal tumors compared to normal gastric tissue (Figure 21B).

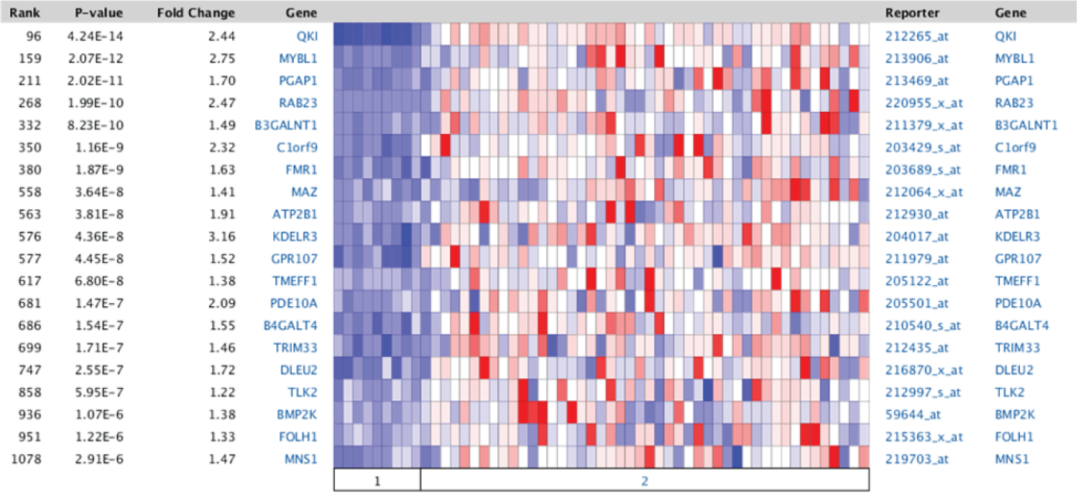
A

Barretina Sarcoma

Nat Genet 2010/07/04 158 samples

mRNA 12,624 measured genes

Human Genome U133A Array



Rank	P-value	Fold Change	Gene	Reporter	Gene
96	4.24E-14	2.44	QKI	212265_at	QKI
159	2.07E-12	2.75	MYBL1	213906_at	MYBL1
211	2.02E-11	1.70	PGAP1	213469_at	PGAP1
268	1.99E-10	2.47	RAB23	220955_x_at	RAB23
332	8.23E-10	1.49	B3GALNT1	211379_x_at	B3GALNT1
350	1.16E-9	2.32	C1orf9	203429_s_at	C1orf9
380	1.87E-9	1.63	FMR1	203689_s_at	FMR1
558	3.64E-8	1.41	MAZ	212064_x_at	MAZ
563	3.81E-8	1.91	ATP2B1	212930_at	ATP2B1
576	4.36E-8	3.16	KDELR3	204017_at	KDELR3
577	4.45E-8	1.52	GPR107	211979_at	GPR107
617	6.80E-8	1.38	TMEFF1	205122_at	TMEFF1
681	1.47E-7	2.09	PDE10A	205501_at	PDE10A
686	1.54E-7	1.55	B4GALT4	210540_s_at	B4GALT4
699	1.71E-7	1.46	TRIM33	212435_at	TRIM33
747	2.55E-7	1.72	DLEU2	216870_x_at	DLEU2
858	5.95E-7	1.22	TLK2	212997_s_at	TLK2
936	1.07E-6	1.38	BMP2K	59644_at	BMP2K
951	1.22E-6	1.33	FOLH1	215363_x_at	FOLH1
1078	2.91E-6	1.47	MNS1	219703_at	MNS1

Legend

1. Adipose Tissue (9) 2. Dedifferentiated Liposarcoma (46)

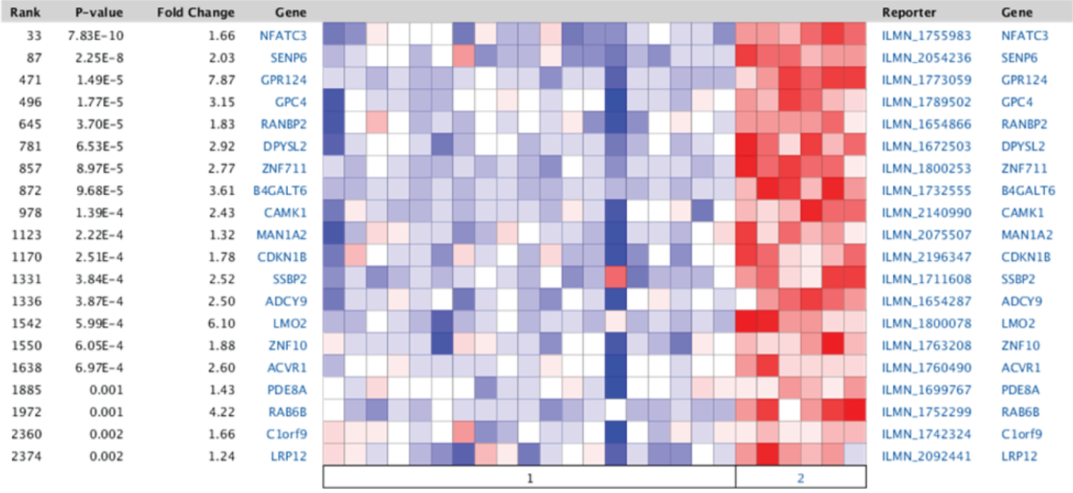
B

Cho Gastric

Clin Cancer Res 2011/04/01 90 samples

mRNA 19,273 measured genes

Illumina HumanWG-6 v3.0 Expression Beadchip



Rank	P-value	Fold Change	Gene	Reporter	Gene
33	7.83E-10	1.66	NFATC3	ILMN_1755983	NFATC3
87	2.25E-8	2.03	SENP6	ILMN_2054236	SENP6
471	1.49E-5	7.87	GPR124	ILMN_1773059	GPR124
496	1.77E-5	3.15	GPC4	ILMN_1789502	GPC4
645	3.70E-5	1.83	RANBP2	ILMN_1654866	RANBP2
781	6.53E-5	2.92	DPYSL2	ILMN_1672503	DPYSL2
857	8.97E-5	2.77	ZNF711	ILMN_1800253	ZNF711
872	9.68E-5	3.61	B4GALT6	ILMN_1732555	B4GALT6
978	1.39E-4	2.43	CAMK1	ILMN_2140990	CAMK1
1123	2.22E-4	1.32	MAN1A2	ILMN_2075507	MAN1A2
1170	2.51E-4	1.78	CDKN1B	ILMN_2196347	CDKN1B
1331	3.84E-4	2.52	SSBP2	ILMN_1711608	SSBP2
1336	3.87E-4	2.50	ADCY9	ILMN_1654287	ADCY9
1542	5.99E-4	6.10	LMO2	ILMN_1800078	LMO2
1550	6.05E-4	1.88	ZNF10	ILMN_1763208	ZNF10
1638	6.97E-4	2.60	ACVR1	ILMN_1760490	ACVR1
1885	0.001	1.43	PDE8A	ILMN_1699767	PDE8A
1972	0.001	4.22	RAB6B	ILMN_1752299	RAB6B
2360	0.002	1.66	C1orf9	ILMN_1742324	C1orf9
2374	0.002	1.24	LRP12	ILMN_2092441	LRP12

Legend

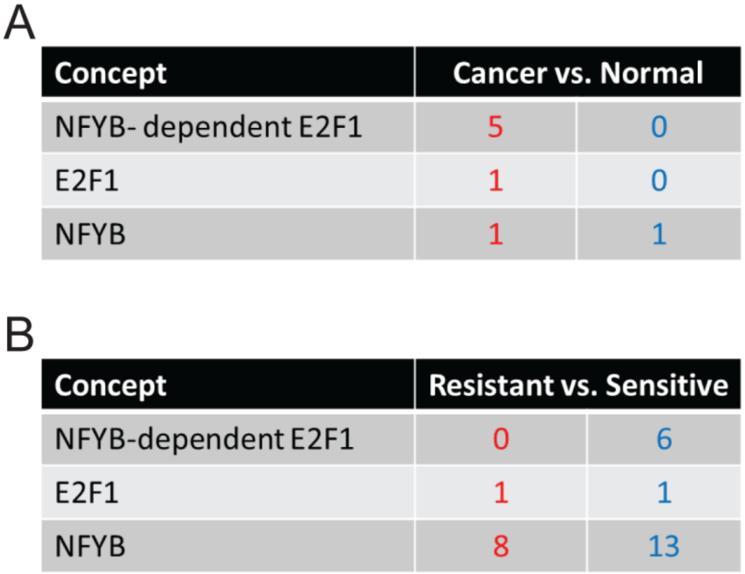
1. Gastric Tissue (19) 2. Gastrointestinal Stromal Tumor (6)

Figure 21. NFYB signature overexpression associates with sarcoma

A) Oncomine analysis demonstrating upregulation of genes from the NFYB-dependent E2F1 signature in Barretina sarocoma dataset. Heatmap showing higher expression in dedifferentiated liposarcoma compared to normal adipose tissue. Red color designates high expression and blue color designates low expression. B) Oncomine analysis demonstrating upregulation of genes from the NFYB-dependent E2F1 signature in gastrointestinal stromal tumor compared to normal gastric tissue.

To determine whether this association is specific for the E2F1/NFYB signature, we compared the results to those signatures generated using the Oncomine determined concept "Up-regulated genes in human osteosarcoma cell line (U2OS) expressing E2F1 compared to U2OS controls – Literature-defined Concepts" (E2F1 signature). This signature did not show the same enrichment within sarcoma datasets, with only one dataset in which the signature significantly associates with a concept overexpressed in cancer compared to normal (Figure 22A). Similarly, an E2F1-independent NFYB signature in which the top 145 genes that were most strongly downregulated by NFYB knockdown showed only one dataset in which the signature significantly associates with a concept overexpressed in cancer and one in which it associates with a concept underexpressed in cancer. Together with the data showing that NFYB is protective against E2F1-mediated apoptosis, these results suggest that the NFYB-dependent E2F1 transcriptional program we defined may be involved in sarcomagenesis or maintenance.

The Oncomine analysis further indicated that the E2F1/NFY signature is significantly associated with concepts from six independent datasets showing lower expression of the genes in drug sensitive sarcoma cell lines compared with drug resistant lines (Figure 22B). This association is not limited to sarcoma, but persists across a broad range of cell lines. Two example heatmaps show overexpression of signature genes in irinotecan resistant cell lines (Figure 23A) and panobinostat resistant cell lines (Figure 23B) compared to sensitive cell lines. We again explored whether this association was specific to the joint E2F1/NFY signature by comparing it to similar analysis with the NFYB-independent E2F1 signature and E2F1-independent NFYB signature. Unlike the joint signature, the separate E2F1 and NFYB signatures did not show a clear association with drug resistance (Figure 22).



A

Concept	Cancer vs. Normal	
NFYB- dependent E2F1	5	0
E2F1	1	0
NFYB	1	1

B

Concept	Resistant vs. Sensitive	
NFYB-dependent E2F1	0	6
E2F1	1	1
NFYB	8	13

Figure 22. Signature association with disease state and therapy resistance

A) Number of sarcoma datasets from cancer vs normal Oncomine analysis in which signatures (concepts) consisting of differentially expressed genes are significantly associated with the indicated signatures. The NFYB-dependent E2F1 signature (Figure 2D and Supplemental Table 2) is derived from our microarray analysis and described in the Results section. The E2F1 (U2OS induced) signature is a concept defined by Oncomine based on previously published microarray analysis of genes upregulated in U2OS ER-E2F1 cells following OHT induction. The E2F1-independent NFYB signature consists of probes (174) which expression is reduced 30% following NFYB knockdown regardless of the effect of E2F1 induction. Significance is set at an odds ratio of at least two and p value of less than 0.0001. Red indicates number of signatures overexpressed in cancer compared to normal cells. Blue represents number of signatures under-expressed in cancer compared to normal. B) Number of datasets in which concepts from therapy sensitivity datasets are significantly associated with the signatures described in (A). Significance is set at an odds ratio of at least two and p value of less than 0.0001. Red indicates number of signatures overexpressed in drug resistant compared to drug sensitive cells. Blue represents number of signatures under-expressed in resistant when compared to sensitive cells.

Barretina CellLine

Nature 2012/03/28 917 samples
mRNA 19,574 measured genes
Human Genome U133 Plus 2.0 Array

Rank	P-value	Fold Change	Gene	Reporter	Gene
14	3.51E-12	-6.24	KDELR3	204017_at	KDELR3
120	1.46E-9	-2.96	COBLL1	203642_s_at	COBLL1
176	6.05E-9	-4.16	JAG1	216268_s_at	JAG1
209	1.57E-8	-1.78	B4GALT4	210540_s_at	B4GALT4
238	2.79E-8	-2.58	ADCY9	204497_at	ADCY9
276	5.26E-8	-2.65	ACVR1	203935_at	ACVR1
373	2.64E-7	-4.20	TGFA	205016_at	TGFA
423	4.67E-7	-3.50	MYO6	203215_s_at	MYO6
599	2.86E-6	-5.08	CLDN1	222549_at	CLDN1
895	2.56E-5	-2.75	ITGB4	204990_s_at	ITGB4
938	3.18E-5	-2.31	GPC4	204984_at	GPC4
1035	5.72E-5	-2.71	IL1R1	202948_at	IL1R1
1061	6.50E-5	-1.29	GPR107	211979_at	GPR107
1305	1.77E-4	-1.82	MYO1E	203072_at	MYO1E
1307	1.78E-4	-1.39	EPB41L4A	220119_at	EPB41L4A
1607	4.83E-4	-2.16	DPP4	203716_s_at	DPP4
1836	9.01E-4	-1.21	FAM63B	37802_r_at	FAM63B
2073	0.002	-1.31	GLUD2	210447_at	GLUD2
2075	0.002	-1.62	PRKCZ	202178_at	PRKCZ
2102	0.002	-1.39	RAB23	223463_at	RAB23

Legend

1. Irinotecan Resistant (84) 2. Irinotecan Sensitive (71)

Barretina CellLine

Nature 2012/03/28 917 samples
mRNA 19,574 measured genes
Human Genome U133 Plus 2.0 Array

Rank	P-value	Fold Change	Gene	Reporter	Gene
66	3.66E-19	-11.37	CLDN1	222549_at	CLDN1
324	2.33E-13	-5.80	KDELR3	204017_at	KDELR3
542	2.81E-11	-4.49	TGFA	205016_at	TGFA
654	1.99E-10	-2.70	ACVR1	203935_at	ACVR1
682	2.98E-10	-3.39	ITGB4	204990_s_at	ITGB4
686	3.10E-10	-2.49	COBLL1	203642_s_at	COBLL1
710	4.13E-10	-3.80	IL1R1	202948_at	IL1R1
874	3.14E-9	-3.34	JAG1	216268_s_at	JAG1
958	8.90E-9	-1.61	B4GALT4	210540_s_at	B4GALT4
1023	1.79E-8	-2.27	MYO1E	203072_at	MYO1E
1137	5.40E-8	-1.57	B4GALT1	201883_s_at	B4GALT1
1209	9.33E-8	-1.58	PDE8A	212522_at	PDE8A
1223	1.00E-7	-2.27	ADCY9	204497_at	ADCY9
1327	2.10E-7	-2.43	PTHLH	211756_at	PTHLH
1425	4.27E-7	-1.62	RAB23	223463_at	RAB23
1495	7.04E-7	-2.60	MYO6	203216_s_at	MYO6
1894	8.68E-6	-2.33	PTGS2	204748_at	PTGS2
1984	1.40E-5	-1.84	HPCAL1	205462_s_at	HPCAL1
1993	1.46E-5	-1.28	GPR107	211979_at	GPR107
2102	2.61E-5	-1.72	MNS1	219703_at	MNS1

Legend

1. Panobinostat Resistant (164) 2. Panobinostat Sensitive (86)

Figure 23. NFYB signature overexpression associates with drug therapy resistance

A) Oncomine analysis demonstrating upregulation of genes from the NFYB-dependent E2F1 signature in in irinotecan resistant cell lines compared to irinotecan sensitive cell lines. B) Oncomine analysis demonstrating upregulation of genes from the NFYB-dependent E2F1 signature in panobinostat resistant cell lines compared to panobinostat sensitive cell lines.

In summary, when considered together with data showing NFYB as protective against E2F1-mediated apoptosis, this analysis suggests that the joint E2F1/NFYB transcriptional program may play a role in cancer resistance to chemotherapy.

3.4 Discussion

E2Fs play a critical role in many cell functions including proliferation, development, and cell death. Among the E2Fs, E2F1 is of particular interest because it activates expression of genes involved in both proliferation and apoptosis (208), with studies detailing transcription factor partners as helping to determine specificity towards one or the other (128). Our study identifies NFYB as a novel direct E2F1 target that, in turn, regulates multiple downstream genes. The regulomes of E2F1 and NFYB partially overlap and in the current study we focused on a group of genes that are induced by both transcription factors. Based on previously published studies we suggest that some of these joint targets are responsible for the attenuation of E2F1-induced apoptosis (129, 203), thus suggesting a complex interplay regulating the final apoptotic phenotype. Furthermore, in silico analysis of this subset of co-regulated genes suggests that the NFYB-dependent E2F1 program is linked with sarcoma as it is significantly associated with signatures upregulated in various types of sarcoma compared to normal tissue. This analysis furthermore demonstrates that E2F1/NFY signature is associated with chemotherapy resistance.

In addition to genes induced by both E2F1 and NFYB our clustering analysis also identified a group of genes that are reduced by E2F1 in an NFYB-dependent manner (Fig 16A, cluster 5). These include multiple mitotic genes, such as cdc25C as well as the other two subunits of the NFY complex, NFYA and NFYC. These results may indicate a compensatory induction of NFYA and NFYC in response to the depletion of NFYB. Additional studies are needed to examine the functional significance of this complex interplay between E2F1 and different subunits of the NFY complex in cell cycle progression and other biological processes.

A recent study has shown that E2F1 is transcriptionally induced by ectopic NFYA expression and that NFYA binds to the E2F1 promoter (198). Combined with our results showing E2F1 induces NFYB by direct binding to its promoter indicates a positive feedback loop in which the two transcription factors amplify one another's expression. Notably, however, whereas NFYA was shown to promote apoptosis by inducing E2F1, our results suggest a pro-survival function for NFYB. Such opposing effects of the two subunits of the NF-Y complex might reveal different aspects of co-regulation and feedback mechanisms between E2F1 and NF-Y or be a result of different approaches employed (ectopic expression of NFYA in Gurtner et al vs knockdown of the endogenous NFYB in our study) or different cell lines used.

Our results indicated a number of cell cycle genes regulated by NFYB. These include CCNB1, CCNB2, CDC25C, CDK2, NEK2, PLK1, AKAP8, and GMNN. Given the abundance of CCAAT boxes in the promoters of genes regulated during the G_2/M phase, this is not surprising. Cyclin B1, Cyclin B, and CDC25C have been shown to be direct targets of NF-Y and that this binding changes dynamically through the different phases of the cell cycle (192, 193, 209-211). Furthermore, CDK2 is known to phosphorylate NFYA and physically associate with NF-Y and that this phosphorylation is essential for DNA binding of NF-Y (212) as well as the expression of its target cell cycle regulatory genes (213). The microarray data indicates that CCNB1, CCNB2, and CDK2 are regulated by NF-Y given that its downregulated following knockdown of NFYB. Interestingly, CDC25C expression increases following NFYB knockdown, suggesting NF-Y negatively regulates the gene's transcription.

Our study also identified several pro- and anti-apoptotic genes regulated by NF-Y and E2F1. API5, an apoptosis inhibitor, is down-regulated while SIVA1, an apoptosis-inducing factor, is upregulated following NFYB knockdown. This result seems to match phenotypically with our result that cleaved PARP levels increase upon NFYB depletion. However, expression of TIMP3 and TNFRSF10B, both pro-apoptotic, also decrease upon NFYB knockdown, suggesting a balance between pro and anti-apoptotic factors that under different conditions may change NFYB's role in E2F1-mediated apoptosis. This mirrors the results of studies indicating that NF-Y plays roles in proliferation and survival as well as cell death. Further studies to determine the effect that levels of pro and anti-apoptotic targets regulated by NFYB and E2F1 have on E2F1-mediated apoptosis are necessary to determine the key players of the distinct life or death fates in different contexts.

This study also identified an E2F1 and NFYB co-regulated gene, SFRP1, which may mediate upregulation of β-catenin pathway activity following NFYB knockdown through inhibition of Wnt. The effect of this regulation can be seen in the components of the Wnt/ β-catenin pathway that appear within the co-regulated gene set. The pathway may also play an important role in the protective effect of NFYB against E2F1-mediated apoptosis given its importance in survival and proliferation (214). However, additional studies are needed to determine whether SFRP1 is a direct target of NFYB, whether knockdown of SFRP1 directly contributes to β-catenin pathway upregulation, and whether SFRP1 mRNA or protein levels inversely correlate with survival as implied by the data.

Finally, we identify a set of genes regulated by E2F1 and NFYB that is overexpressed in sarcoma compared to normal tissues. This is intriguing in that the inducible E2F1 system used in this study comes from an osteosarcoma cell line, indicating that the NFYB-dependent E2F1 transcriptional program may be specific to tissues that have a mesenchymal origin. Further enhancing this idea is the fact that E2F1 target genes induced in an NFYB-independent manner do not show this association with sarcoma. Our results showing NFYB to be protective against E2F1 mediated apoptosis moreover suggest this NFYB/E2F1 regulated gene set may be involved in the development or survival of sarcoma cells. In addition to its association with sarcoma, a subset of NFYB-dependent E2F1 regulated genes is overexpressed in chemotherapy resistant cell lines. This is not limited to a narrow range of chemotherapies, as the association is found for both a topoisomerase inhibitor (irinotecan) and an HDAC inhibitor (panobinostat). Intriguingly, this association does not seem to be limited to sarcoma. Unfortunately, the clinical relevance of these associations is limited by the lack of publicly available sarcoma datasets with patient survival data. Once such data becomes available, potential use of the NFYB/E2F signature can be explored as a biomarker for selecting patients that may respond to certain chemotherapies in future clinical trials.

4. Conclusions

4.1 Summary

This thesis describes the utilization of DNA microarrays to study heterogeneity in the tumors' response to targeted therapies and oncogenesis. First, we demonstrated Eμ-Myc tumors can be accurately and reproducibly classified into two subgroups. Importantly, these two subgroups displayed differences in surface cell staining, oncogenic and tumor suppressor protein expression, and copy number changes. We also showed that the two subgroups also differed significantly based on their predicted signatures for cell of origin and treatment response. By utilizing oncogenic pathway signatures, we explored the segregation of patterns of pathway activation between the two groups and showed their similarity to gene expression patterns seen in human aggressive B cell lymphoma. Finally, we were able to take advantage of the predicted pathway status for TNFα to show that pathway scores can predict the likely response to a targeted therapy for that specific pathway.

In studying oncogenic mechanisms, we made use of an inducible E2F1 system that recapitulates ectopic E2F1 expression to induce apoptosis. By examining genome-wide expression following E2F1 activation in this system, we were able to identify NFYB as a novel and direct target of E2F1 via chromatin immunoprecipitation. Furthermore, we were able to identify an NFYB-dependent E2F1 transcriptional program that included a number of pro- and antiapoptotic genes, with SFRP1, API5 and SIVA1 being most interesting in terms of further experimentation. In terms of its direct effect on E2F1-mediated apoptosis, we showed that NFYB is protective against apoptosis induced by ectopic E2F1 expression. We also identified an NFYB-dependent E2F1 signature that, based on *in silico* analysis, is overexpressed in sarcoma, tumors that derive from cells of mesenchymal origin. Combined with the fact that NFYB protects cells from E2F1 induced apoptosis, these data point to the potential roles of the joint E2F1/NFYB transcriptional program as being protective against sarcomagenesis and maintenance.

4.2 Limitations and future directions

4.2.1 Eμ-Myc heterogeneity

With every study, there are limitations present that should be addressed in future experiments. The heterogeneity present within the Eμ-Myc model is apparent, but what drives the heterogeneity is still not well understood. For example, our results indicated that Myc protein levels are also different between the two clusters, with cluster one tumors having uniformly higher Myc protein relative to cluster two. Given the importance of the Myc transcription factor in cellular processes, this segregation could play a significant role in cluster differences. How much this plays into the differences in patterns of gene expression will need to be empirically determined. A good place to start would be to knock down the expression of Myc in cluster one tumors and look at the resulting changes in patterns of pathway activation and other gene expression characteristics.

Traditional treatment of human lymphoma is a multi-agent regimen administered over the course of several weeks. The mouse chemotherapy studies presented here are single agent single dose administrations that may not reflect the combinatorial efficacy seen in human disease. If a safe dosage in mice can be established for using multiple agents over several administrations, then these study results should be replicated in that context. In the case of heterogeneity in treatment response to targeted therapies, the two lymphoma subgroups showed segregation in other oncogenic pathways. STAT3, PI3K, and E2F1 all have small molecule drugs targeting different pathway components. For example, in the case of E2F1, cyclin dependent kinase inhibitors, which prevent E2F activation, would be predicted to have greater efficacy in cluster one tumors. Many targeted therapies are also used in combination with traditional chemotherapy with enhanced combinatorial effect. This is something that can also be tested in the mouse model. We initiated work towards this goal by combining cyclophosphamide or doxorubicin with bortezomib, but initial studies were limited by death of the mice within a day of drug administration with mice showing symptoms similar to tumor lysis syndrome. The confirmation of tumor lysis through elevated blood phosphate levels compared to untreated tumor bearing mice indicates the possible enhanced efficacy of combining a chemotherapy and targeted therapy.

Our analysis of the relationship between p53 mutational status and chemotherapy resistance also had limitations in the fact that we were only able to test one cluster one tumor with wild type p53. Additional tumors of this type will need to tested with a rigorous multivariate analysis. Given that the Western analysis of the p53 axis proteins showed possibly aberrant expression of ARF and Mdm2 in several cluster one and two tumors, the p53 sequencing analysis should be expanded to encompass the p53 axis, which includes sequencing ARF and MDM2, as these are also frequently mutated. With the decreasing costs of sequencing, it would be incredibly informative to conduct transcriptome sequencing of Eμ-Myc tumors as this would not only provide expression information, but also show the breadth of mutations that have occurred in the expressed portion of the genome. One of the limitations of transcriptome sequencing is its inability to show copy number changes. Approaches such as exome sequencing or whole genome sequencing would complement RNA sequencing to provide information on non-transcribed mutations that may affect gene expression as well as copy number alterations, which are commonly seen in all types of cancers. All the aforementioned sequencing methods applied over dozens of Eμ-Myc tumors may indicate driver mutations through the enrichment of mutations in certain genes. Sequencing data would also provide an opportunity to see just how well the mouse model recapitulates the mutational landscape of BL and DLBCL.

4.2.2 NFY in E2F1-mediated apoptosis

In terms of the relationship between E2F1 and NFYB, I believe we have only scratched the surface. I was able to show that several co-regulated genes were direct targets of E2F1, but was unable to do the same for NFY. Given the importance of NFYA in recognizing the CCAAT promoter sequence, using NFYB as the immunoprecipitating protein may cause many target genes to not be detected. Chromatin immunoprecipitation experiments with a tagged NFYA may resolve this issue. An alternative strategy would be to use a luciferase assay in which the target gene promoters are fused with the luciferase gene. By assaying luciferase intensity following NFYB ectopic expression or knockdown may also indicate these genes are direct NF-Y targets. Further experiments are needed to see which co-regulated targets are most responsible for the protective effect of NFYB against E2F1-mediated apoptosis. Knockdown experiments of induced genes would begin to give a better picture of which components of E2F1-induced apoptosis are moderated by NF-Y. Additionally, it would be interesting to see what role p53 plays, as it is involved in regulation of both pathways. Despite the lack of drugs targeting NF-Y components, one can imagine that with greater understanding of how E2F1 and NF-Y determine their involvement in proliferation or apoptosis, it will eventually be possible for drugs to target the proliferative potential of the two transcription factors while leaving the apoptotic potential intact.

The co-regulation of SFRP1, a well-known Wnt/β catenin inhibitor, is interesting because of the results from pathway activation analyses showing that β catenin pathway activity is predicted to increase following NFYB and hence SFRP1 mRNA knockdown.

The effect seems to be short-lived, however, as this difference is only significant at zero and four hours, but results at the eight hour timepoint are close to significance. Experiments in which SFRP1 levels are knocked down independently of NFYB are needed to solidify the link between SFRP1 and β catenin in this cell line. Also assaying for apoptosis under modification of SFRP1 levels would indicate whether or not the β catenin pathway plays a role in E2F1 mediated apoptosis.

Furthermore, there is currently a separation between the NF-Y regulome which contains apoptotic genes such as API5 and SIVA1 and the NFYB dependent E2F1 transcriptional program which is linked to sarcoma and drug resistance. The protective effect of NFYB in E2F1-mediated apoptosis is the only phenotype linking them. Other phenotypic assays such as Caspase Glo or flow cytometry for caspase can confirm the cleaved PARP Western results. The role that the apoptotic genes play in both E2F1-induced apoptosis as well as its enhancement upon knockdown of NFYB can also link the two sets of genes. With the NFYB overexpression cell line, it would also be feasible to test if overexpression of NFY attenuates E2F1-mediated apoptosis. Initial testing based on cleaved PARP Western blotting has indicated that the effect is negligible, but alternative assays should be done to confirm this result.

The Oncomine analysis indicated a connection between the NFYB-dependent E2F1 transcriptional program and sarcoma. Interestingly, the U2OS that this inducible E2F1 system is based on is derived from an osteosarcoma. It may be because of this that the NFYB-dependent E2F1 concept shows any association at all with sarcoma. To establish whether this gene set plays a role in other cancers, it may be necessary to recreate the inducible E2F1 system in another cell line and show that inducing E2F1 in such a context leads to apoptosis. However, the fact that the coregulated gene set associates with sarcoma indicates that the genes may play an important role in tumorigenesis of or tumor maintenance in mesenchymal cells. An Oncomine analysis should also be run on the NF-Y regulome described in the microarray analysis as this list of genes is more extensive than the NFYB concept used in the Oncomine analysis.

In summary, the work presented in this thesis provides a view of cancer heterogeneity and oncogenic mechanisms as recognized through the prism of patterns of gene expression that can be used to extrapolate meaningful information for directing targeted therapeutics. It is part of the work of developing tools that can make personalized medicine more effective. By combining genomic analysis of cancers with functional studies of relevant biological pathways to understand the mechanisms behind their development, we can begin to harness the potential of individualized medicine.

Appendix A

Table 9: Genes from hierarchical clustering of NFYB regulated genes

Cluster 1

Probe.Set.ID	Gene Title
205359_at	A kinase (PRKA) anchor protein 6
210355_at	parathyroid hormone-like hormone
202948_at	interleukin 1 receptor, type I
215363_x_at	folate hydrolase (prostate-specific membrane antigen) 1
214023_x_at	tubulin, beta 2B
212959_s_at	N-acetylglucosamine-1-phosphate transferase, alpha and beta subunits
221761_at	adenylosuccinate synthase
222071_s_at	solute carrier organic anion transporter family, member 4C1
203072_at	myosin IE
209925_at	occludin pseudogene; occludin
207362_at	solute carrier family 30 (zinc transporter), member 4
215446_s_at	lysyl oxidase
204062_s_at	unc-51-like kinase 2 (C. elegans)
202037_s_at	secreted frizzled-related protein 1
202036_s_at	secreted frizzled-related protein 1
201625_s_at	insulin induced gene 1
211478_s_at	dipeptidyl-peptidase 4
216256_at	glutamate receptor, metabotropic 8
216255_s_at	glutamate receptor, metabotropic 8
215245_x_at	fragile X mental retardation 1
203689_s_at	fragile X mental retardation 1
219526_at	chromosome 14 open reading frame 169
214578_s_at	similar to Rho-associated, coiled-coil containing protein kinase 1; Rho-associated, coiled-coil containing protein kinase 1
216205_s_at	mitofusin 2
202460_s_at	lipin 2
206233_at	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 6
218129_s_at	nuclear transcription factor Y, beta
218128_at	nuclear transcription factor Y, beta
218127_at	nuclear transcription factor Y, beta
206615_s_at	ADAM metallopeptidase domain 22
204720_s_at	DnaJ (Hsp40) homolog, subfamily C, member 6

219703_at	meiosis-specific nuclear structural 1
202178_at	protein kinase C, zeta
203429_s_at	chromosome 1 open reading frame 9
211130_x_at	ectodysplasin A
210829_s_at	single-stranded DNA binding protein 2
218223_s_at	pleckstrin homology domain containing, family O member 1
210941_at	protocadherin 7
202842_s_at	DnaJ (Hsp40) homolog, subfamily B, member 9
204479_at	osteoclast stimulating factor 1
212870_at	son of sevenless homolog 2 (Drosophila)
209569_x_at	DNA segment on chromosome 4 (unique) 234 expressed sequence
204160_s_at	ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function)
219932_at	solute carrier family 27 (fatty acid transporter), member 6
208964_s_at	fatty acid desaturase 1
208962_s_at	fatty acid desaturase 1
208963_x_at	fatty acid desaturase 1
216167_at	leucine rich repeat neuronal 2
208925_at	claudin domain containing 1
209993_at	ATP-binding cassette, sub-family B (MDR/TAP), member 1
212812_at	serine incorporator 5

Cluster 2

Probe.Set.ID	Gene Title
211048_s_at	protein disulfide isomerase family A, member 4
218095_s_at	transmembrane protein 165
215489_x_at	homer homolog 3 (Drosophila)
208703_s_at	amyloid beta (A4) precursor-like protein 2
202101_s_at	v-ral simian leukemia viral oncogene homolog B (ras related; GTP binding protein)
221318_at	neurogenic differentiation 4
203548_s_at	lipoprotein lipase
202516_s_at	discs, large homolog 1 (Drosophila)
217761_at	acireductone dioxygenase 1
217931_at	canopy 3 homolog (zebrafish)
204298_s_at	lysyl oxidase
211804_s_at	cyclin-dependent kinase 2
201761_at	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase

201662_s_at	acyl-CoA synthetase long-chain family member 3
209706_at	NK3 homeobox 1
201063_at	reticulocalbin 1, EF-hand calcium binding domain
203688_at	polycystic kidney disease 2 (autosomal dominant)
201490_s_at	peptidylprolyl isomerase F
219628_at	zinc finger, matrin type 3
213369_at	protocadherin 21
206683_at	zinc finger protein 165
200927_s_at	RAB14, member RAS oncogene family
214959_s_at	API5-like 1; apoptosis inhibitor 5
205573_s_at	sorting nexin 7
206491_s_at	N-ethylmaleimide-sensitive factor attachment protein, alpha
208309_s_at	mucosa associated lymphoid tissue lymphoma translocation gene 1
221781_s_at	DnaJ (Hsp40) homolog, subfamily C, member 10
209232_s_at	dynactin 5 (p25)

Cluster 3

Probe.Set.ID	Gene Title
219730_at	mediator complex subunit 18
214642_x_at	melanoma antigen family A, 5
203179_at	galactose-1-phosphate uridylyltransferase
215076_s_at	collagen, type III, alpha 1
201852_x_at	collagen, type III, alpha 1
206463_s_at	dehydrogenase/reductase (SDR family) member 2
214079_at	dehydrogenase/reductase (SDR family) member 2
209208_at	mannose-P-dolichol utilization defect 1
204256_at	ELOVL family member 6, elongation of long chain fatty acids (FEN1/Elo2, SUR4/Elo3-like, yeast)
210868_s_at	ELOVL family member 6, elongation of long chain fatty acids (FEN1/Elo2, SUR4/Elo3-like, yeast)
201399_s_at	translocation associated membrane protein 1
211162_x_at	stearoyl-CoA desaturase (delta-9-desaturase)
202061_s_at	sel-1 suppressor of lin-12-like (C. elegans)
202062_s_at	sel-1 suppressor of lin-12-like (C. elegans)
218396_at	vacuolar protein sorting 13 homolog C (S. cerevisiae)
209295_at	tumor necrosis factor receptor superfamily, member 10b
203827_at	WD repeat domain, phosphoinositide interacting 1
210970_s_at	inhibitor of Bruton agammaglobulinemia tyrosine kinase

209234_at	kinesin family member 1B
220205_at	transmembrane phosphatase with tensin homology
209198_s_at	synaptotagmin XI
202816_s_at	synovial sarcoma translocation, chromosome 18
201860_s_at	plasminogen activator, tissue
205523_at	hyaluronan and proteoglycan link protein 1
212403_at	ubiquitin protein ligase E3B
209550_at	necdin homolog (mouse)
221730_at	collagen, type V, alpha 2
211080_s_at	NIMA (never in mitosis gene a)-related kinase 2
204641_at	NIMA (never in mitosis gene a)-related kinase 2
219049_at	chondroitin sulfate N-acetylgalactosaminyltransferase 1
208861_s_at	alpha thalassemia/mental retardation syndrome X-linked (RAD54 homolog, S. cerevisiae)
204149_s_at	glutathione S-transferase mu 4
201932_at	leucine rich repeat containing 41
210406_s_at	RAB6C, member RAS oncogene family; RAB6A, member RAS oncogene family; hypothetical LOC100130819; RAB6C-like
218539_at	F-box protein 34
217188_s_at	chromosome 14 open reading frame 1
207714_s_at	serpin peptidase inhibitor, clade H (heat shock protein 47), member 1, (collagen binding protein 1)
203069_at	synaptic vesicle glycoprotein 2A
217749_at	coatomer protein complex, subunit gamma
214710_s_at	cyclin B1
209118_s_at	tubulin, alpha 1a
206369_s_at	phosphoinositide-3-kinase, catalytic, gamma polypeptide
203738_at	chromosome 5 open reading frame 22
218124_at	retinol saturase (all-trans-retinol 13,14-reductase)
203211_s_at	myotubularin related protein 2
212613_at	butyrophilin, subfamily 3, member A2
209846_s_at	butyrophilin, subfamily 3, member A2
211063_s_at	NCK adaptor protein 1
204725_s_at	NCK adaptor protein 1
202043_s_at	spermine synthase; similar to spermine synthase
202705_at	cyclin B2
208180_s_at	histone cluster 1, H4l; histone cluster 1, H4k; histone cluster 4, H4; histone cluster 1, H4h; histone cluster 1, H4j; histone cluster 1, H4i; histone cluster 1, H4d; histone cluster 1, H4c; histone cluster 1, H4f; histone cluster 1, H4e; histone cluster 1, H4b; histone cluster 1, H4a; histone cluster 2, H4a; histone

	cluster 2, H4b
203441_s_at	cadherin 2, type 1, N-cadherin (neuronal)
217977_at	selenoprotein X, 1
209079_x_at	protocadherin gamma subfamily C, 3; protocadherin gamma subfamily C, 5; protocadherin gamma subfamily C, 4; protocadherin gamma subfamily A, 12
211066_x_at	protocadherin gamma subfamily C, 3; protocadherin gamma subfamily C, 5; protocadherin gamma subfamily C, 4; protocadherin gamma subfamily A, 12
215836_s_at	protocadherin gamma subfamily C, 3; protocadherin gamma subfamily C, 5; protocadherin gamma subfamily C, 4; protocadherin gamma subfamily A, 12
205717_x_at	protocadherin gamma subfamily C, 3; protocadherin gamma subfamily C, 5; protocadherin gamma subfamily C, 4; protocadherin gamma subfamily A, 12
211959_at	insulin-like growth factor binding protein 5
212274_at	lipin 1
222209_s_at	transmembrane protein 135
214945_at	family with sequence similarity 153, member B
201633_s_at	cytochrome b5 type B (outer mitochondrial membrane)
202920_at	ankyrin 2, neuronal
204749_at	nucleosome assembly protein 1-like 3
218047_at	oxysterol binding protein-like 9
213645_at	enolase superfamily member 1
201147_s_at	TIMP metallopeptidase inhibitor 3
205822_s_at	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1 (soluble)
201341_at	ectodermal-neural cortex (with BTB-like domain)
203041_s_at	lysosomal-associated membrane protein 2
218150_at	ADP-ribosylation factor-like 5A

Cluster 4

Probe.Set.ID	Gene Title
220097_s_at	transmembrane protein 104
209478_at	stimulated by retinoic acid 13 homolog (mouse)
213577_at	squalene epoxidase
203489_at	SIVA1, apoptosis-inducing factor
210792_x_at	SIVA1, apoptosis-inducing factor
213147_at	homeobox A10
208366_at	protocadherin 11 X-linked
220983_s_at	sprouty homolog 4 (Drosophila)

213089_at	ENSG00000219982
201565_s_at	inhibitor of DNA binding 2, dominant negative helix-loop-helix protein
218642_s_at	coiled-coil-helix-coiled-coil-helix domain containing 7
218350_s_at	geminin, DNA replication inhibitor
204790_at	SMAD family member 7
201419_at	BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase)
204621_s_at	nuclear receptor subfamily 4, group A, member 2
215599_at	glucuronidase, beta pseudogene
215043_s_at	glucuronidase, beta pseudogene
206565_x_at	glucuronidase, beta pseudogene
204770_at	transporter 2, ATP-binding cassette, sub-family B (MDR/TAP)

Cluster 5

Probe.Set.ID	Gene Title
221777_at	chromosome 12 open reading frame 52
218609_s_at	nudix (nucleoside diphosphate linked moiety X)-type motif 2
209223_at	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 2, 8kDa
209682_at	Cas-Br-M (murine) ecotropic retroviral transforming sequence b
215498_s_at	mitogen-activated protein kinase kinase 3
215499_at	mitogen-activated protein kinase kinase 3
203709_at	phosphorylase kinase, gamma 2 (testis)
204218_at	chromosome 11 open reading frame 51
203848_at	A kinase (PRKA) anchor protein 8
203719_at	excision repair cross-complementing rodent repair deficiency, complementation group 1 (includes overlapping antisense sequence)
202215_s_at	nuclear transcription factor Y, gamma
205828_at	matrix metallopeptidase 3 (stromelysin 1, progelatinase)
40569_at	myeloid zinc finger 1
203931_s_at	mitochondrial ribosomal protein L12
32259_at	enhancer of zeste homolog 1 (Drosophila)
206785_s_at	killer cell lectin-like receptor subfamily C, member 1
203478_at	NADH dehydrogenase (ubiquinone) 1, subcomplex unknown, 1, 6kDa
218597_s_at	CDGSH iron sulfur domain 1
216852_x_at	immunoglobulin lambda variable 2-11; immunoglobulin lambda constant 2 (Kern-Oz- marker); immunoglobulin lambda variable 1-44; immunoglobulin lambda constant 1 (Mcg marker); immunoglobulin lambda variable 1-40; immunoglobulin lambda variable 3-21; immunoglobulin lambda locus; immunoglobulin lambda constant 3 (Kern-Oz+ marker)

202736_s_at	LSM4 homolog, U6 small nuclear RNA associated (S. cerevisiae)
37462_i_at	splicing factor 3a, subunit 2, 66kDa
202240_at	polo-like kinase 1 (Drosophila)
219053_s_at	vacuolar protein sorting 37 homolog C (S. cerevisiae)
213897_s_at	mitochondrial ribosomal protein L23
214110_s_at	lymphocyte-specific protein 1 pseudogene
218741_at	centromere protein M
215380_s_at	gamma-glutamyl cyclotransferase
218774_at	decapping enzyme, scavenger
215293_s_at	post-GPI attachment to proteins 2
41858_at	post-GPI attachment to proteins 2
220089_at	L-2-hydroxyglutarate dehydrogenase
218866_s_at	polymerase (RNA) III (DNA directed) polypeptide K, 12.3 kDa
204475_at	matrix metallopeptidase 1 (interstitial collagenase)
205167_s_at	cell division cycle 25 homolog C (S. pombe)
217010_s_at	cell division cycle 25 homolog C (S. pombe)
204107_at	nuclear transcription factor Y, alpha
204108_at	nuclear transcription factor Y, alpha
204109_s_at	nuclear transcription factor Y, alpha
33304_at	interferon stimulated exonuclease gene 20kDa
217907_at	mitochondrial ribosomal protein L18
205730_s_at	actin binding LIM protein family, member 3
203880_at	COX17 cytochrome c oxidase assembly homolog (S. cerevisiae)
201774_s_at	non-SMC condensin I complex, subunit D2
215734_at	chromosome 19 open reading frame 36
213893_x_at	PMS2 postmeiotic segregation increased 2 (S. cerevisiae)-like
214756_x_at	postmeiotic segregation increased 2-like 1 pseudogene
217485_x_at	postmeiotic segregation increased 2-like 1 pseudogene
203980_at	fatty acid binding protein 4, adipocyte
214073_at	cortactin
65517_at	adaptor-related protein complex 1, mu 2 subunit
218261_at	adaptor-related protein complex 1, mu 2 subunit
213322_at	chromosome 6 open reading frame 130

Appendix B

Table 10: Genes coregulated by E2F1 and NFYB

Probe.Set.ID	Gene Title
200762_at	dihydropyrimidinase-like 2
201283_s_at	trafficking protein, kinesin binding 1
201688_s_at	tumor protein D52
201689_s_at	tumor protein D52
201693_s_at	early growth response 1
201711_x_at	RAN binding protein 2
202035_s_at	secreted frizzled-related protein 1
202036_s_at	secreted frizzled-related protein 1
202037_s_at	secreted frizzled-related protein 1
202178_at	protein kinase C, zeta
202236_s_at	solute carrier family 16 (monocarboxylic acid transporters), member 1
202583_s_at	RAN binding protein 9
202670_at	mitogen-activated protein kinase kinase 1
202842_s_at	DnaJ (Hsp40) homolog, subfamily B, member 9
202843_at	ralA binding protein 1
202948_at	interleukin 1 receptor, type I
203044_at	carbohydrate (chondroitin) synthase 1
203057_s_at	PR domain containing 2, with ZNF domain
203072_at	myosin IE
203311_s_at	ADP-ribosylation factor 6
203520_s_at	zinc finger protein 318
203625_x_at	S-phase kinase-associated protein 2 (p45)
203641_s_at	COBL-like 1
203689_s_at	fragile X mental retardation 1
203935_at	activin A receptor, type I
204062_s_at	unc-51-like kinase 2 (C. elegans)
204145_at	FSHD region gene 1
204160_s_at	ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function)
204161_s_at	ectonucleotide pyrophosphatase/phosphodiesterase 4 (putative function)
204184_s_at	adrenergic, beta, receptor kinase 2
204194_at	BTB and CNC homology 1, basic leucine zipper transcription factor 1
204249_s_at	LIM domain only 2 (rhombotin-like 1)
204392_at	calcium/calmodulin-dependent protein kinase I

204479_at	osteoclast stimulating factor 1
204497_at	adenylate cyclase 9
204507_s_at	protein phosphatase 3 (formerly 2B), regulatory subunit B, 19kDa, alpha isoform (calcineurin B, type I)
204526_s_at	TBC1 domain family, member 8 (with GRAM domain)
204720_s_at	DnaJ (Hsp40) homolog, subfamily C, member 6
204748_at	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)
204832_s_at	bone morphogenetic protein receptor, type IA
204953_at	synaptosomal-associated protein, 91kDa homolog (mouse)
204984_at	glypican 4
204989_s_at	integrin, beta 4
205015_s_at	transforming growth factor, alpha
205016_at	transforming growth factor, alpha
205123_s_at	transmembrane protein with EGF-like and two follistatin-like domains 1
205164_at	glycine C-acetyltransferase (2-amino-3-ketobutyrate coenzyme A ligase)
205333_s_at	RCE1 homolog, prenyl protein peptidase (S. cerevisiae)
205462_s_at	hippocalcin-like 1
205478_at	protein phosphatase 1, regulatory (inhibitor) subunit 1A
205789_at	CD1d molecule /// CD1d molecule
205842_s_at	Janus kinase 2 (a protein tyrosine kinase)
205850_s_at	gamma-aminobutyric acid (GABA) A receptor, beta 3
205860_x_at	folate hydrolase (prostate-specific membrane antigen) 1
206159_at	growth differentiation factor 10
206176_at	bone morphogenetic protein 6
206233_at	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 6
206456_at	gamma-aminobutyric acid (GABA) A receptor, alpha 5
206511_s_at	sine oculis homeobox homolog 2 (Drosophila)
206615_s_at	ADAM metallopeptidase domain 22
207150_at	solute carrier family 18 (vesicular acetylcholine), member 3
207265_s_at	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3
207275_s_at	acyl-CoA synthetase long-chain family member 1
207362_at	solute carrier family 30 (zinc transporter), member 4
207767_s_at	early growth response 4
207768_at	early growth response 4
207781_s_at	zinc finger protein 711
207824_s_at	MYC-associated zinc finger protein (purine-binding transcription factor)
208237_x_at	ADAM metallopeptidase domain 22

208244_at	bone morphogenetic protein 3 (osteogenic)
208264_s_at	eukaryotic translation initiation factor 3, subunit 1 alpha, 35kDa
208606_s_at	wingless-type MMTV integration site family, member 4 /// wingless-type MMTV integration site family, member 4
208652_at	protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform
208985_s_at	eukaryotic translation initiation factor 3, subunit 1 alpha, 35kDa
208990_s_at	heterogeneous nuclear ribonucleoprotein H3 (2H9)
209098_s_at	jagged 1 (Alagille syndrome)
209112_at	cyclin-dependent kinase inhibitor 1B (p27, Kip1)
209281_s_at	ATPase, Ca++ transporting, plasma membrane 1
209339_at	seven in absentia homolog 2 (Drosophila) /// seven in absentia homolog 2 (Drosophila)
209347_s_at	v-maf musculoaponeurotic fibrosarcoma oncogene homolog (avian)
209569_x_at	DNA segment on chromosome 4 (unique) 234 expressed sequence
209570_s_at	DNA segment on chromosome 4 (unique) 234 expressed sequence
209590_at	Bone morphogenetic protein 7 (osteogenic protein 1)
209591_s_at	bone morphogenetic protein 7 (osteogenic protein 1)
209990_s_at	gamma-aminobutyric acid (GABA) B receptor, 2
210021_s_at	uracil-DNA glycosylase 2
210127_at	RAB6B, member RAS oncogene family
210190_at	syntaxin 11
210240_s_at	cyclin-dependent kinase inhibitor 2D (p19, inhibits CDK4)
210355_at	parathyroid hormone-like hormone
210447_at	glutamate dehydrogenase 2
210480_s_at	myosin VI
210540_s_at	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 4
210555_s_at	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 3
210716_s_at	restin (Reed-Steinberg cell-expressed intermediate filament-associated protein)
210829_s_at	single-stranded DNA binding protein 2
210875_s_at	transcription factor 8 (represses interleukin 2 expression)
211067_s_at	growth arrest-specific 7 /// growth arrest-specific 7
211171_s_at	phosphodiesterase 10A
211379_x_at	UDP-GalNAc:betaGlcNAc beta 1,3-galactosaminyltransferase, polypeptide 1 (Globoside blood group)
211478_s_at	dipeptidyl-peptidase 4 (CD26, adenosine deaminase complexing protein 2)
211631_x_at	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 1
211812_s_at	UDP-GalNAc:betaGlcNAc beta 1,3-galactosaminyltransferase, polypeptide 1 (Globoside blood group)

211985_s_at	calmodulin 1 (phosphorylase kinase, delta)
212056_at	KIAA0182
212209_at	thyroid hormone receptor associated protein 2
212435_at	tripartite motif-containing 33
212447_at	kelch repeat and BTB (POZ) domain containing 2
212521_s_at	phosphodiesterase 8A
212750_at	protein phosphatase 1, regulatory (inhibitor) subunit 16B
212812_at	Serine incorporator 5
212870_at	Ras association (RalGDS/AF-6) domain family 3
212930_at	ATPase, Ca++ transporting, plasma membrane 1
212986_s_at	tousled-like kinase 2
213353_at	ATP-binding cassette, sub-family A (ABC1), member 5
213469_at	GPI deacylase
213470_s_at	heterogeneous nuclear ribonucleoprotein H1 (H)
213533_at	DNA segment on chromosome 4 (unique) 234 expressed sequence
213695_at	paraoxonase 3
213906_at	v-myb myeloblastosis viral oncogene homolog (avian)-like 1
214449_s_at	ras homolog gene family, member Q
214543_x_at	quaking homolog, KH domain RNA binding (mouse)
214578_s_at	similar to Rho-associated protein kinase 1 (Rho-associated, coiled-coil containing protein kinase 1) (p160 ROCK-1) (p160ROCK)
214691_x_at	family with sequence similarity 63, member B
214790_at	SUMO1/sentrin specific peptidase 6
214890_s_at	DKFZP564J102 protein
214954_at	sushi domain containing 5
215245_x_at	fragile X mental retardation 1
215363_x_at	folate hydrolase (prostate-specific membrane antigen) 1
215716_s_at	ATPase, Ca++ transporting, plasma membrane 1
215794_x_at	glutamate dehydrogenase 2
216125_s_at	RAN binding protein 9
216255_s_at	glutamate receptor, metabotropic 8
216256_at	glutamate receptor, metabotropic 8
216350_s_at	zinc finger protein 10
216488_s_at	ATPase, Class VI, type 11A
216521_s_at	BRCA1/BRCA2-containing complex, subunit 3
216627_s_at	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 1
216870_x_at	deleted in lymphocytic leukemia, 2
216953_s_at	Wilms tumor 1

217280_x_at	gamma-aminobutyric acid (GABA) A receptor, alpha 5
217644_s_at	son of sevenless homolog 2 (Drosophila)
217920_at	mannosidase, alpha, class 1A, member 2
218127_at	nuclear transcription factor Y, beta
218128_at	nuclear transcription factor Y, beta
218129_s_at	nuclear transcription factor Y, beta
218182_s_at	claudin 1
218223_s_at	pleckstrin homology domain containing, family O member 1
218319_at	pellino homolog 1 (Drosophila)
218499_at	Mst3 and SOK1-related kinase
219312_s_at	zinc finger and BTB domain containing 10
219631_at	low density lipoprotein-related protein 12
219703_at	meiosis-specific nuclear structural 1
219778_at	zinc finger protein, multitype 2
219797_at	mannosyl (alpha-1,3-)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase, isozyme A
219864_s_at	Down syndrome critical region gene 1-like 2
219892_at	transmembrane 6 superfamily member 1
219932_at	solute carrier family 27 (fatty acid transporter), member 6
220014_at	mesenchymal stem cell protein DSC54
220120_s_at	erythrocyte membrane protein band 4.1 like 4A
220253_s_at	low density lipoprotein-related protein 12
220254_at	low density lipoprotein-related protein 12
220265_at	G protein-coupled receptor 107
220386_s_at	echinoderm microtubule associated protein like 4
220955_x_at	RAB23, member RAS oncogene family
221039_s_at	development and differentiation enhancing factor 1
221268_s_at	sphingosine-1-phosphate phosphatase 1 /// sphingosine-1-phosphate phosphatase 1
221428_s_at	transducin (beta)-like 1X-linked receptor 1 /// transducin (beta)-like 1X-linked receptor 1
221814_at	G protein-coupled receptor 124
222071_s_at	solute carrier organic anion transporter family, member 4C1
222121_at	Src homology 3 domain-containing guanine nucleotide exchange factor
222235_s_at	chondroitin sulfate GalNAcT-2
41577_at	protein phosphatase 1, regulatory (inhibitor) subunit 16B
59644_at	BMP2 inducible kinase

References

1. Hanahan D & Weinberg RA (2000) The hallmarks of cancer. *Cell* 100(1):57-70.

2. Hanahan D & Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* 144(5):646-674.

3. Lander ES, *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* 409(6822):860-921.

4. Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of molecular biology* 98(3):503-517.

5. Augenlicht LH & Kobrin D (1982) Cloning and screening of sequences expressed in a mouse colon tumor. *Cancer Res* 42(3):1088-1093.

6. Augenlicht LH, *et al.* (1987) Expression of cloned sequences in biopsies of human colonic tissue and in colonic carcinoma cells induced to differentiate in vitro. *Cancer Res* 47(22):6017-6021.

7. Augenlicht LH, Taylor J, Anderson L, & Lipkin M (1991) Patterns of gene expression that characterize the colonic mucosa in patients at genetic risk for colonic cancer. *Proc Natl Acad Sci U S A* 88(8):3286-3289.

8. Schena M, Shalon D, Davis RW, & Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270(5235):467-470.

9. Golub TR, *et al.* (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286(5439):531-537.

10. Alizadeh AA, *et al.* (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403(6769):503-511.

11. Perou CM, *et al.* (2000) Molecular portraits of human breast tumours. *Nature* 406(6797):747-752.

12. Garber ME, *et al.* (2001) Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci U S A* 98(24):13784-13789.

13. Bhattacharjee A, *et al.* (2001) Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A* 98(24):13790-13795.

14. Chaib H, Cockrell EK, Rubin MA, & Macoska JA (2001) Profiling and verification of gene expression patterns in normal and malignant human prostate tissues by cDNA microarray analysis. *Neoplasia* 3(1):43-52.

15. Ramaswamy S, Ross KN, Lander ES, & Golub TR (2003) A molecular signature of metastasis in primary solid tumors. *Nat Genet* 33(1):49-54.

16. van 't Veer LJ, *et al.* (2003) Expression profiling predicts outcome in breast cancer. *Breast cancer research : BCR* 5(1):57-58.

17. van de Vijver MJ, *et al.* (2002) A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 347(25):1999-2009.

18. Lamb J, *et al.* (2003) A mechanism of cyclin D1 action encoded in the patterns of gene expression in human cancer. *Cell* 114(3):323-334.

19. Huang E, *et al.* (2003) Gene expression phenotypic models that predict the activity of oncogenic pathways. *Nat Genet* 34(2):226-230.

20. Bild AH, *et al.* (2006) Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 439(7074):353-357.

21. Gatza ML, *et al.* (2010) A pathway-based classification of human breast cancer. *Proc Natl Acad Sci U S A* 107(15):6994-6999.

22. Chi JT, *et al.* (2006) Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. *PLoS Med* 3(3):e47.

23. Erler JT, *et al.* (2006) Lysyl oxidase is essential for hypoxia-induced metastasis. *Nature* 440(7088):1222-1226.

24. Nuyten DS, *et al.* (2006) Predicting a local recurrence after breast-conserving therapy by gene expression profiling. *Breast Cancer Res* 8(5):R62.

25. Nuyten DS, Hastie T, Chi JT, Chang HY, & van de Vijver MJ (2008) Combining biological gene expression signatures in predicting outcome in breast cancer: An alternative to supervised classification. *Eur J Cancer* 44(15):2319-2329.

26. Chan DA, *et al.* (2009) Tumor vasculature is regulated by PHD2-mediated angiogenesis and bone marrow-derived cell recruitment. *Cancer Cell* 15(6):527-538.

27. Chen JL, *et al.* (2008) The genomic analysis of lactic acidosis and acidosis response in human cancers. *PLoS Genet* 4(12):e1000293.

28. Chen JL, *et al.* (2010) Lactic acidosis triggers starvation response with paradoxical induction of TXNIP through MondoA. *PLoS Genet* 6(9).

29. Sheth SS, *et al.* (2006) Hepatocellular carcinoma in Txnip-deficient mice. *Oncogene* 25(25):3528-3536.

30. Muoio DM (2007) TXNIP links redox circuitry to glucose control. *Cell Metab* 5(6):412-414.

31. Parikh H, *et al.* (2007) TXNIP regulates peripheral glucose metabolism in humans. *PLoS Med* 4(5):e158.

32. Hui ST, *et al.* (2008) Txnip balances metabolic and growth signaling via PTEN disulfide reduction. *Proc Natl Acad Sci U S A* 105(10):3921-3926.

33. Gatza ML, *et al.* (2011) Analysis of tumor environmental response and oncogenic pathway activation identifies distinct basal and luminal features in HER2-related breast tumor subtypes. *Breast Cancer Res* 13(3):R62.

34. Kung HN, Marks JR, & Chi JT (2011) Glutamine synthetase as genetic determinant of cell type specific glutamine independency in breast epithelia. *PLoS Genet* In Press.

35. Lucas JE, Kung HN, & Chi JT (2010) Latent factor analysis to discover pathway-associated putative segmental aneuploidies in human cancers. *PLoS Comput Biol* 6(9):e1000920.

36. Weinstein IB (2002) Cancer. Addiction to oncogenes--the Achilles heal of cancer. *Science* 297(5578):63-64.

37. Melo JV (1996) The molecular biology of chronic myeloid leukaemia. *Leukemia* 10(5):751-756.

38. Stuttfeld E & Ballmer-Hofer K (2009) Structure and function of VEGF receptors. *IUBMB life* 61(9):915-922.

39. Levitzki A (2013) Tyrosine kinase inhibitors: views of selectivity, sensitivity, and clinical performance. *Annual review of pharmacology and toxicology* 53:161-185.

40. Takeuchi K, *et al.* (2012) RET, ROS1 and ALK fusions in lung cancer. *Nat Med* 18(3):378-381.

41. Hay N & Sonenberg N (2004) Upstream and downstream of mTOR. *Genes Dev* 18(16):1926-1945.

42. Davies H, *et al.* (2002) Mutations of the BRAF gene in human cancer. *Nature* 417(6892):949-954.

43. Kohler G & Milstein C (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256(5517):495-497.

44. Stern M & Herrmann R (2005) Overview of monoclonal antibodies in cancer therapy: present and promise. *Critical reviews in oncology/hematology* 54(1):11-29.

45. Chothia C, *et al.* (1989) Conformations of immunoglobulin hypervariable regions. *Nature* 342(6252):877-883.

46. Chapman K, *et al.* (2009) Preclinical development of monoclonal antibodies: considerations for the use of non-human primates. *mAbs* 1(5):505-516.

47. Perantoni AO, Rice JM, Reed CD, Watatani M, & Wenk ML (1987) Activated neu oncogene sequences in primary tumors of the peripheral nervous system induced in rats by transplacental exposure to ethylnitrosourea. *Proc Natl Acad Sci U S A* 84(17):6317-6321.

48. Drebin JA, Link VC, Weinberg RA, & Greene MI (1986) Inhibition of tumor growth by a monoclonal antibody reactive with an oncogene-encoded tumor antigen. *Proc Natl Acad Sci U S A* 83(23):9129-9133.

49. Drebin JA, Link VC, Stern DF, Weinberg RA, & Greene MI (1986) Development of monoclonal antibodies reactive with the product of the neu oncogene. *Symposium on Fundamental Cancer Research* 38:277-289.

50. Mitri Z, Constantine T, & O'Regan R (2012) The HER2 Receptor in Breast Cancer: Pathophysiology, Clinical Use, and New Advances in Therapy. *Chemotherapy research and practice* 2012:743193.

51. Cragg MS, Walshe CA, Ivanov AO, & Glennie MJ (2005) The biology of CD20 and its potential as a target for mAb therapy. *Current directions in autoimmunity* 8:140-174.

52. Sassoon I & Blanc V (2013) Antibody-drug conjugate (ADC) clinical pipeline: a review. *Methods in molecular biology* 1045:1-27.

53. Sheiness D & Bishop JM (1979) DNA and RNA from uninfected vertebrate cells contain nucleotide sequences related to the putative transforming gene of avian myelocytomatosis virus. *Journal of virology* 31(2):514-521.

54. Hu SS, Lai MM, & Vogt PK (1979) Genome of avian myelocytomatosis virus MC29: analysis by heteroduplex mapping. *Proc Natl Acad Sci U S A* 76(3):1265-1268.

55. Brooks TA & Hurley LH (2010) Targeting MYC Expression through G-Quadruplexes. *Genes & cancer* 1(6):641-649.

56. Hurley LH, Von Hoff DD, Siddiqui-Jain A, & Yang D (2006) Drug targeting of the c-MYC promoter to repress gene expression via a G-quadruplex silencer element. *Seminars in oncology* 33(4):498-512.

57. Sears R, *et al.* (2000) Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability. *Genes Dev* 14(19):2501-2514.

58. Dalla-Favera R, *et al.* (1982) Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc Natl Acad Sci U S A* 79(24):7824-7827.

59. Taub R, *et al.* (1982) Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. *Proc Natl Acad Sci U S A* 79(24):7837-7841.

60. Chen Y, *et al.* (2014) Identification of druggable cancer driver genes amplified across TCGA datasets. *PloS one* 9(5):e98293.

61. Ross JS, *et al.* (2013) Comprehensive genomic profiling of epithelial ovarian cancer by next generation sequencing-based diagnostic assay reveals new routes to targeted therapies. *Gynecologic oncology* 130(3):554-559.

62. Nesbit CE, Tersak JM, & Prochownik EV (1999) MYC oncogenes and human neoplastic disease. *Oncogene* 18(19):3004-3016.

63. Eischen CM, Weber JD, Roussel MF, Sherr CJ, & Cleveland JL (1999) Disruption of the ARF-Mdm2-p53 tumor suppressor pathway in Myc-induced lymphomagenesis. *Genes Dev* 13(20):2658-2669.

64. Eischen CM, *et al.* (2001) Bcl-2 is an apoptotic target suppressed by both c-Myc and E2F-1. *Oncogene* 20(48):6983-6993.

65. Maclean KH, Keller UB, Rodriguez-Galindo C, Nilsson JA, & Cleveland JL (2003) c-Myc augments gamma irradiation-induced apoptosis by suppressing Bcl-XL. *Molecular and cellular biology* 23(20):7256-7270.

66. Stewart TA, Pattengale PK, & Leder P (1984) Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. *Cell* 38(3):627-637.

67. Hundley JE, *et al.* (1997) Differential regulation of cell cycle characteristics and apoptosis in MMTV-myc and MMTV-ras mouse mammary tumors. *Cancer Res* 57(4):600-603.

68. Sinn E, *et al.* (1987) Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: synergistic action of oncogenes in vivo. *Cell* 49(4):465-475.

69. Andrechek ER, *et al.* (2009) Genetic heterogeneity of Myc-induced mammary tumors reflecting diverse phenotypes including metastatic potential. *Proc Natl Acad Sci U S A* 106(38):16387-16392.

70. Adams JM, *et al.* (1985) The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. *Nature* 318(6046):533-538.

71. Harris AW, *et al.* (1988) The E mu-myc transgenic mouse. A model for high-incidence spontaneous lymphoma and leukemia of early B cells. *J Exp Med* 167(2):353-371.

72. Alt JR, Greiner TC, Cleveland JL, & Eischen CM (2003) Mdm2 haplo-insufficiency profoundly inhibits Myc-induced lymphomagenesis. *The EMBO journal* 22(6):1442-1450.

73. Eischen CM, Roussel MF, Korsmeyer SJ, & Cleveland JL (2001) Bax loss impairs Myc-induced apoptosis and circumvents the selection of p53 mutations during Myc-mediated lymphomagenesis. *Molecular and cellular biology* 21(22):7653-7662.

74. Strasser A, Harris AW, Bath ML, & Cory S (1990) Novel primitive lymphoid tumours induced in transgenic mice by cooperation between myc and bcl-2. *Nature* 348(6299):331-333.

75. Tao W & Levine AJ (1999) P19(ARF) stabilizes p53 by blocking nucleo-cytoplasmic shuttling of Mdm2. *Proc Natl Acad Sci U S A* 96(12):6937-6941.

76. Weber JD, Taylor LJ, Roussel MF, Sherr CJ, & Bar-Sagi D (1999) Nucleolar Arf sequesters Mdm2 and activates p53. *Nature cell biology* 1(1):20-26.

77. Momand J, Zambetti GP, Olson DC, George D, & Levine AJ (1992) The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* 69(7):1237-1245.

78. Honda R & Yasuda H (1999) Association of p19(ARF) with Mdm2 inhibits ubiquitin ligase activity of Mdm2 for tumor suppressor p53. *The EMBO journal* 18(1):22-27.

79. Roth J, Dobbelstein M, Freedman DA, Shenk T, & Levine AJ (1998) Nucleo-cytoplasmic shuttling of the hdm2 oncoprotein regulates the levels of the p53 protein via a pathway used by the human immunodeficiency virus rev protein. *The EMBO journal* 17(2):554-564.

80. Barak Y, Juven T, Haffner R, & Oren M (1993) mdm2 expression is induced by wild type p53 activity. *The EMBO journal* 12(2):461-468.

81. Wu X, Bayle JH, Olson D, & Levine AJ (1993) The p53-mdm-2 autoregulatory feedback loop. *Genes Dev* 7(7A):1126-1132.

82. Reed JC, *et al.* (1996) BCL-2 family proteins: regulators of cell death involved in the pathogenesis of cancer and resistance to therapy. *Journal of cellular biochemistry* 60(1):23-32.

83. Schmitt CA, McCurrach ME, de Stanchina E, Wallace-Brodeur RR, & Lowe SW (1999) INK4a/ARF mutations accelerate lymphomagenesis and promote chemoresistance by disabling p53. *Genes Dev* 13(20):2670-2677.

84. Odvody J, *et al.* (2010) A deficiency in Mdm2 binding protein inhibits Myc-induced B-cell proliferation and lymphomagenesis. *Oncogene* 29(22):3287-3296.

85. Ren B, *et al.* (2002) E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. *Genes Dev* 16(2):245-256.

86. Cam H, *et al.* (2004) A common set of gene regulatory networks links metabolism and growth inhibition. *Molecular cell* 16(3):399-411.

87. Bracken AP, Ciro M, Cocito A, & Helin K (2004) E2F target genes: unraveling the biology. *Trends in biochemical sciences* 29(8):409-417.

88. Knudson AG, Jr. (1971) Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* 68(4):820-823.

89. Friend SH, *et al.* (1986) A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* 323(6089):643-646.

90. Schwarz JK, *et al.* (1993) Interactions of the p107 and Rb proteins with E2F during the cell proliferation response. *The EMBO journal* 12(3):1013-1020.

91. Chellappan SP, Hiebert S, Mudryj M, Horowitz JM, & Nevins JR (1991) The E2F transcription factor is a cellular target for the RB protein. *Cell* 65(6):1053-1061.

92. Bagchi S, Weinmann R, & Raychaudhuri P (1991) The retinoblastoma protein copurifies with E2F-I, an E1A-regulated inhibitor of the transcription factor E2F. *Cell* 65(6):1063-1072.

93. Chittenden T, Livingston DM, & Kaelin WG, Jr. (1991) The T/E1A-binding domain of the retinoblastoma product can interact selectively with a sequence-specific DNA-binding protein. *Cell* 65(6):1073-1082.

94. Bandara LR & La Thangue NB (1991) Adenovirus E1a prevents the retinoblastoma gene product from complexing with a cellular transcription factor. *Nature* 351(6326):494-497.

95. Hiebert SW, Chellappan SP, Horowitz JM, & Nevins JR (1992) The interaction of RB with E2F coincides with an inhibition of the transcriptional activity of E2F. *Genes Dev* 6(2):177-185.

96. Zamanian M & La Thangue NB (1992) Adenovirus E1a prevents the retinoblastoma gene product from repressing the activity of a cellular transcription factor. *The EMBO journal* 11(7):2603-2610.

97. Dimova DK & Dyson NJ (2005) The E2F transcriptional network: old acquaintances with new faces. *Oncogene* 24(17):2810-2826.

98. Sherr CJ (1994) G1 phase progression: cycling on cue. *Cell* 79(4):551-555.

99. Sherr CJ & McCormick F (2002) The RB and p53 pathways in cancer. *Cancer cell* 2(2):103-112.

100. Borlado LR & Mendez J (2008) CDC6: from DNA replication to cell cycle checkpoints and oncogenesis. *Carcinogenesis* 29(2):237-243.

101. Hwang HC & Clurman BE (2005) Cyclin E in normal and neoplastic cell cycles. *Oncogene* 24(17):2776-2786.

102. Wu L*, et al.* (2001) The E2F1-3 transcription factors are essential for cellular proliferation. *Nature* 414(6862):457-462.

103. Aslanian A, Iaquinta PJ, Verona R, & Lees JA (2004) Repression of the Arf tumor suppressor by E2F3 is required for normal cell cycle kinetics. *Genes Dev* 18(12):1413-1422.

104. Rowland BD*, et al.* (2002) E2F transcriptional repressor complexes are critical downstream targets of p19(ARF)/p53-induced proliferative arrest. *Cancer cell* 2(1):55-65.

105. Timmers C*, et al.* (2007) E2f1, E2f2, and E2f3 control E2F target expression and cellular proliferation via a p53-dependent negative feedback loop. *Molecular and cellular biology* 27(1):65-78.

106. Muller H*, et al.* (2001) E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. *Genes Dev* 15(3):267-285.

107. Iaquinta PJ & Lees JA (2007) Life and death decisions by the E2F transcription factors. *Current opinion in cell biology* 19(6):649-657.

108. Zhu JW, DeRyckere D, Li FX, Wan YY, & DeGregori J (1999) A role for E2F1 in the induction of ARF, p53, and apoptosis during thymic negative selection. *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research* 10(12):829-838.

109. Rogoff HA, Pickering MT, Debatis ME, Jones S, & Kowalik TF (2002) E2F1 induces phosphorylation of p53 that is coincident with p53 accumulation and apoptosis. *Molecular and cellular biology* 22(15):5308-5318.

110. Powers JT, *et al.* (2004) E2F1 uses the ATM signaling pathway to induce p53 and Chk2 phosphorylation and apoptosis. *Molecular cancer research : MCR* 2(4):203-214.

111. Chen D, Padiernos E, Ding F, Lossos IS, & Lopez CD (2005) Apoptosis-stimulating protein of p53-2 (ASPP2/53BP2L) is an E2F target gene. *Cell death and differentiation* 12(4):358-368.

112. Fogal V, *et al.* (2005) ASPP1 and ASPP2 are new transcriptional targets of E2F. *Cell death and differentiation* 12(4):369-376.

113. Hershko T, Chaussepied M, Oren M, & Ginsberg D (2005) Novel link between E2F and p53: proapoptotic cofactors of p53 are transcriptionally upregulated by E2F. *Cell death and differentiation* 12(4):377-383.

114. Pan H, *et al.* (1998) Key roles for E2F1 in signaling p53-dependent apoptosis and in cell division within developing tumors. *Molecular cell* 2(3):283-292.

115. Polager S & Ginsberg D (2009) p53 and E2f: partners in life and death. *Nat Rev Cancer* 9(10):738-748.

116. Stanelle J & Putzer BM (2006) E2F1-induced apoptosis: turning killers into therapeutics. *Trends in molecular medicine* 12(4):177-185.

117. Pop C, Timmer J, Sperandio S, & Salvesen GS (2006) The apoptosome activates caspase-9 by dimerization. *Molecular cell* 22(2):269-275.

118. Riedl SJ, Li W, Chao Y, Schwarzenbacher R, & Shi Y (2005) Structure of the apoptotic protease-activating factor 1 bound to ADP. *Nature* 434(7035):926-933.

119. Rossi M, *et al.* (2005) The ubiquitin-protein ligase Itch regulates p73 stability. *The EMBO journal* 24(4):836-848.

120. Hershko T & Ginsberg D (2004) Up-regulation of Bcl-2 homology 3 (BH3)-only proteins by E2F1 mediates apoptosis. *The Journal of biological chemistry* 279(10):8627-8634.

121. Dick FA & Dyson N (2003) pRB contains an E2F1-specific binding domain that allows E2F1-induced apoptosis to be regulated separately from other E2F activities. *Molecular cell* 12(3):639-649.

122. Markham D, Munro S, Soloway J, O'Connor DP, & La Thangue NB (2006) DNA-damage-responsive acetylation of pRb regulates binding to E2F-1. *EMBO reports* 7(2):192-198.

123. Pediconi N, *et al.* (2003) Differential regulation of E2F1 apoptotic target genes in response to DNA damage. *Nature cell biology* 5(6):552-558.

124. Lin WC, Lin FT, & Nevins JR (2001) Selective induction of E2F1 in response to DNA damage, mediated by ATM-dependent phosphorylation. *Genes Dev* 15(14):1833-1844.

125. Urist M, Tanaka T, Poyurovsky MV, & Prives C (2004) p73 induction after DNA damage is regulated by checkpoint kinases Chk1 and Chk2. *Genes Dev* 18(24):3041-3054.

126. Hallstrom TC & Nevins JR (2003) Specificity in the activation and control of transcription factor E2F-dependent apoptosis. *Proc Natl Acad Sci U S A* 100(19):10848-10853.

127. Hallstrom TC & Nevins JR (2006) Jab1 is a specificity factor for E2F1-induced apoptosis. *Genes Dev* 20(5):613-623.

128. Shats I, *et al.* (2013) FOXO transcription factors control E2F1 transcriptional specificity and apoptotic function. *Cancer Res* 73(19):6056-6067.

129. Morris EJ, *et al.* (2006) Functional identification of Api5 as a suppressor of E2F-dependent apoptosis in vivo. *PLoS Genet* 2(11):e196.

130. Morgensztern D & McLeod HL (2005) PI3K/Akt/mTOR pathway as a target for cancer therapy. *Anti-cancer drugs* 16(8):797-803.

131. Liu K, Paik JC, Wang B, Lin FT, & Lin WC (2006) Regulation of TopBP1 oligomerization by Akt/PKB for cell survival. *The EMBO journal* 25(20):4795-4807.

132. Hallstrom TC, Mori S, & Nevins JR (2008) An E2F1-dependent gene expression program that determines the balance between proliferation and cell death. *Cancer cell* 13(1):11-22.

133. Jaffe ES & Pittaluga S (2011) Aggressive B-cell lymphomas: a review of new and old entities in the WHO classification. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program* 2011:506-514.

134. Dave SS, *et al.* (2006) Molecular diagnosis of Burkitt's lymphoma. *N Engl J Med* 354(23):2431-2442.

135. Lenz G, *et al.* (2008) Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med* 359(22):2313-2323.

136. Monti S, *et al.* (2005) Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. *Blood* 105(5):1851-1861.

137. Shipp MA, *et al.* (2002) Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nat Med* 8(1):68-74.

138. Love C, *et al.* (2012) The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet* 44(12):1321-1325.

139. Morin RD, *et al.* (2013) Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood* 122(7):1256-1265.

140. Lohr JG, *et al.* (2012) Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A* 109(10):3879-3884.

141. Pasqualucci L, *et al.* (2011) Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet* 43(9):830-837.

142. Zhang J, *et al.* (2013) Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A* 110(4):1398-1403.

143. Mori S, *et al.* (2008) Utilization of pathway signatures to reveal distinct types of B lymphoma in the Emicro-myc model and human diffuse large B-cell lymphoma. *Cancer Res* 68(20):8525-8534.

144. Egle A, Harris AW, Bouillet P, & Cory S (2004) Bim is a suppressor of Myc-induced mouse B cell leukemia. *Proc Natl Acad Sci U S A* 101(16):6164-6169.

145. Jing H, *et al.* (2011) Opposing roles of NF-kappaB in anti-cancer treatment outcome unveiled by cross-species investigations. *Genes Dev* 25(20):2137-2146.

146. Lindemann RK, *et al.* (2007) Analysis of the apoptotic and therapeutic activities of histone deacetylase inhibitors by using a mouse model of B cell lymphoma. *Proc Natl Acad Sci U S A* 104(19):8071-8076.

147. Rempel RE, *et al.* (2009) A role for E2F activities in determining the fate of Myc-induced lymphomagenesis. *PLoS Genet* 5(9):e1000640.

148. Schmitt CA & Lowe SW (2001) Bcl-2 mediates chemoresistance in matched pairs of primary E(mu)-myc lymphomas in vivo. *Blood Cells Mol Dis* 27(1):206-216.

149. Schmitt CA, Rosenthal CT, & Lowe SW (2000) Genetic analysis of chemoresistance in primary murine lymphomas. *Nat Med* 6(9):1029-1035.

150. Wendel HG, *et al.* (2004) Survival signalling by Akt and eIF4E in oncogenesis and cancer therapy. *Nature* 428(6980):332-337.

151. Wendel HG, *et al.* (2006) Determinants of sensitivity and resistance to rapamycin-chemotherapy drug combinations in vivo. *Cancer Res* 66(15):7639-7646.

152. Burgess DJ, *et al.* (2008) Topoisomerase levels determine chemotherapy response in vitro and in vivo. *Proc Natl Acad Sci U S A* 105(26):9053-9058.

153. Johnson WE, Li C, & Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8(1):118-127.

154. Pittman J, *et al.* (2004) Integrated modeling of clinical and gene expression information for personalized prediction of disease outcomes. *Proc Natl Acad Sci U S A* 101(22):8431-8436.

155. Tibshirani R, Hastie T, Narasimhan B, & Chu G (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* 99(10):6567-6572.

156. Chang JT, *et al.* (2011) SIGNATURE: A workbench for gene expression signature analysis. *BMC Bioinformatics* 12(1):443.

157. Green MR, *et al.* (2011) Signatures of murine B-cell development implicate Yy1 as a regulator of the germinal center-specific program. *Proc Natl Acad Sci U S A* 108(7):2873-2878.

158. Klapproth K & Wirth T (2010) Advances in the understanding of MYC-induced lymphomagenesis. *Br J Haematol* 149(4):484-497.

159. Hummel M, *et al.* (2006) A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N Engl J Med* 354(23):2419-2430.

160. Rubio-Moscardo F, *et al.* (2005) Characterization of 8p21.3 chromosomal deletions in B-cell lymphoma: TRAIL-R1 and TRAIL-R2 as candidate dosage-dependent tumor suppressor genes. *Blood* 106(9):3214-3222.

161. Gonzalez-Aguilar A, *et al.* (2012) Recurrent mutations of MYD88 and TBL1XR1 in primary central nervous system lymphomas. *Clin Cancer Res*.

162. Ngo VN, *et al.* (2011) Oncogenically active MYD88 mutations in human lymphoma. *Nature* 470(7332):115-119.

163. Seki R, *et al.* (2010) Prognostic significance of S-phase kinase-associated protein 2 and p27kip1 in patients with diffuse large B-cell lymphoma: effects of rituximab. *Ann Oncol* 21(4):833-841.

164. Forshell LP, *et al.* (2011) The direct Myc target Pim3 cooperates with other Pim kinases in supporting viability of Myc-induced B-cell lymphomas. *Oncotarget* 2(6):448-460.

165. The International Non-Hodgkin's Lymphoma Prognostic Factors Project (1993) A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. *N Engl J Med* 329(14):987-994.

166. Ziepert M, *et al.* (2010) Standard International prognostic index remains a valid predictor of outcome for patients with aggressive CD20+ B-cell lymphoma in the rituximab era. *J Clin Oncol* 28(14):2373-2380.

167. Thomas DA, *et al.* (2006) Chemoimmunotherapy with hyper-CVAD plus rituximab for the treatment of adult Burkitt and Burkitt-type lymphoma or acute lymphoblastic leukemia. *Cancer* 106(7):1569-1580.

168. Klapper W, *et al.* (2008) Structural aberrations affecting the MYC locus indicate a poor prognosis independent of clinical risk factors in diffuse large B-cell lymphomas treated within randomized trials of the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL). *Leukemia* 22(12):2226-2229.

169. Xu-Monette ZY, *et al.* (2012) Mutational profile and prognostic significance of TP53 in diffuse large B-cell lymphoma patients treated with R-CHOP: report from an International DLBCL Rituximab-CHOP Consortium Program Study. *Blood* 120(19):3986-3996.

170. Young KH, *et al.* (2008) Structural profiles of TP53 gene mutations predict clinical outcome in diffuse large B-cell lymphoma: an international collaborative study. *Blood* 112(8):3088-3098.

171. Loboda A, *et al.* (2010) A gene expression signature of RAS pathway dependence predicts response to PI3K and RAS pathway inhibitors and expands the population of RAS pathway activated tumors. *BMC Med Genomics* 3:26.

172. Davis RE, Brown KD, Siebenlist U, & Staudt LM (2001) Constitutive nuclear factor kappaB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells. *J Exp Med* 194(12):1861-1874.

173. Klapproth K, Sander S, Marinkovic D, Baumann B, & Wirth T (2009) The IKK2/NF-{kappa}B pathway suppresses MYC-induced lymphomagenesis. *Blood* 114(12):2448-2458.

174. Dunleavy K, *et al.* (2009) Differential efficacy of bortezomib plus chemotherapy within molecular subtypes of diffuse large B-cell lymphoma. *Blood* 113(24):6069-6076.

175. Ruan J, *et al.* (2011) Bortezomib plus CHOP-rituximab for previously untreated diffuse large B-cell lymphoma and mantle cell lymphoma. *J Clin Oncol* 29(6):690-697.

176. Wada Y, *et al.* (2009) A wave of nascent transcription on activated human genes. *Proc Natl Acad Sci U S A* 106(43):18357-18361.

177. Hideshima T, *et al.* (2001) The proteasome inhibitor PS-341 inhibits growth, induces apoptosis, and overcomes drug resistance in human multiple myeloma cells. *Cancer Res* 61(7):3071-3076.

178. Bosco EE & Knudsen ES (2007) RB in breast cancer: at the crossroads of tumorigenesis and treatment. *Cell cycle* 6(6):667-671.

179. Johnson DG & Degregori J (2006) Putting the Oncogenic and Tumor Suppressive Activities of E2F into Context. *Current molecular medicine* 6(7):731-738.

180. Knudsen ES & Knudsen KE (2006) Retinoblastoma tumor suppressor: where cancer meets the cell cycle. *Experimental biology and medicine* 231(7):1271-1281.

181. Ginsberg D (2002) E2F1 pathways to apoptosis. *FEBS letters* 529(1):122-125.

182. Kowalik TF, DeGregori J, Schwarz JK, & Nevins JR (1995) E2F1 overexpression in quiescent fibroblasts leads to induction of cellular DNA synthesis and apoptosis. *Journal of virology* 69(4):2491-2500.

183. Field SJ, *et al.* (1996) E2F-1 functions in mice to promote apoptosis and suppress proliferation. *Cell* 85(4):549-561.

184. Cassimere EK, Pyndiah S, & Sakamuro D (2009) The c-MYC-interacting proapoptotic tumor suppressor BIN1 is a transcriptional target for E2F1 in response to DNA damage. *Cell death and differentiation* 16(12):1641-1653.

185. Croxton R, Ma Y, Song L, Haura EB, & Cress WD (2002) Direct repression of the Mcl-1 promoter by E2F1. *Oncogene* 21(9):1359-1369.

186. Irwin M, *et al.* (2000) Role for the p53 homologue p73 in E2F-1-induced apoptosis. *Nature* 407(6804):645-648.

187. Nahle Z, *et al.* (2002) Direct coupling of the cell cycle and cell death machinery by E2F. *Nature cell biology* 4(11):859-864.

188. Liu K, Luo Y, Lin FT, & Lin WC (2004) TopBP1 recruits Brg1/Brm to repress E2F1-induced apoptosis, a novel pRb-independent and E2F1-specific control for cell survival. *Genes Dev* 18(6):673-686.

189. Elkon R, Linhart C, Sharan R, Shamir R, & Shiloh Y (2003) Genome-wide in silico identification of transcriptional regulators controlling the cell cycle in human cells. *Genome research* 13(5):773-780.

190. Imbriano C, *et al.* (2005) Direct p53 transcriptional repression: in vivo analysis of CCAAT-containing G2/M promoters. *Molecular and cellular biology* 25(9):3737-3751.

191. Romier C, Cocchiarella F, Mantovani R, & Moras D (2003) The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. *The Journal of biological chemistry* 278(2):1336-1345.

192. Bolognese F, *et al.* (1999) The cyclin B2 promoter depends on NF-Y, a trimer whose CCAAT-binding activity is cell-cycle regulated. *Oncogene* 18(10):1845-1853.

193. Farina A, *et al.* (1999) Down-regulation of cyclin B1 gene transcription in terminally differentiated skeletal muscle cells is associated with loss of functional CCAAT-binding NF-Y complex. *Oncogene* 18(18):2818-2827.

194. Korner K, Jerome V, Schmidt T, & Muller R (2001) Cell cycle regulation of the murine cdc25B promoter: essential role for nuclear factor-Y and a proximal repressor element. *The Journal of biological chemistry* 276(13):9662-9669.

195. Zwicker J, *et al.* (1995) Cell cycle regulation of the cyclin A, cdc25C and cdc2 genes is based on a common mechanism of transcriptional repression. *The EMBO journal* 14(18):4514-4522.

196. Benatti P, *et al.* (2008) A balance between NF-Y and p53 governs the pro- and anti-apoptotic transcriptional response. *Nucleic acids research* 36(5):1415-1428.

197. Dolfini D, Zambelli F, Pavesi G, & Mantovani R (2009) A perspective of promoter architecture from the CCAAT box. *Cell cycle* 8(24):4127-4137.

198. Gurtner A, *et al.* (2010) Transcription factor NF-Y induces apoptosis in cells expressing wild-type p53 through E2F1 upregulation and p53 activation. *Cancer Res* 70(23):9711-9720.

199. Giangrande PH, Hallstrom TC, Tunyaplin C, Calame K, & Nevins JR (2003) Identification of E-box factor TFE3 as a functional partner for the E2F3 transcription factor. *Molecular and cellular biology* 23(11):3707-3720.

200. Schlisio S, Halperin T, Vidal M, & Nevins JR (2002) Interaction of YY1 with E2Fs, mediated by RYBP, provides a mechanism for specificity of E2F function. *The EMBO journal* 21(21):5775-5786.

201. Lu H & Hallstrom TC (2013) The nuclear protein UHRF2 is a direct target of the transcription factor E2F1 in the induction of apoptosis. *The Journal of biological chemistry* 288(33):23833-23843.

202. Gatta R, Dolfini D, & Mantovani R (2011) NF-Y joins E2Fs, p53 and other stress transcription factors at the apoptosis table. *Cell death & disease* 2:e162.

203. Qiao Q, *et al.* (2013) Structural architecture of the CARMA1/Bcl10/MALT1 signalosome: nucleation-induced filamentous assembly. *Molecular cell* 51(6):766-779.

204. Prasad KV, *et al.* (1997) CD27, a member of the tumor necrosis factor receptor family, induces apoptosis and binds to Siva, a proapoptotic protein. *Proc Natl Acad Sci U S A* 94(12):6346-6351.

205. Fortin A, *et al.* (2004) The proapoptotic gene SIVA is a direct transcriptional target for the tumor suppressors p53 and E2F1. *The Journal of biological chemistry* 279(27):28706-28714.

206. Finch PW, *et al.* (1997) Purification and molecular cloning of a secreted, Frizzled-related antagonist of Wnt action. *Proc Natl Acad Sci U S A* 94(13):6770-6775.

207. Rhodes DR, *et al.* (2007) Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia* 9(2):166-180.

208. DeGregori J & Johnson DG (2006) Distinct and Overlapping Roles for E2F Family Members in Transcription, Proliferation and Apoptosis. *Current molecular medicine* 6(7):739-748.

209. Manni I, *et al.* (2001) NF-Y mediates the transcriptional inhibition of the cyclin B1, cyclin B2, and cdc25C promoters upon induced G2 arrest. *The Journal of biological chemistry* 276(8):5570-5576.

210. Salsi V, *et al.* (2003) Interactions between p300 and multiple NF-Y trimers govern cyclin B2 promoter function. *The Journal of biological chemistry* 278(9):6642-6650.

211. Sciortino S, *et al.* (2001) The cyclin B1 gene is actively transcribed during mitosis in HeLa cells. *EMBO reports* 2(11):1018-1023.

212. Yun J, *et al.* (2003) Cdk2-dependent phosphorylation of the NF-Y transcription factor and its involvement in the p53-p21 signaling pathway. *The Journal of biological chemistry* 278(38):36966-36972.

213. Chae HD, Yun J, Bang YJ, & Shin DY (2004) Cdk2-dependent phosphorylation of the NF-Y transcription factor is essential for the expression of the cell cycle-regulatory genes and cell cycle G1/S and G2/M transitions. *Oncogene* 23(23):4084-4088.

214. Clevers H & Nusse R (2012) Wnt/beta-catenin signaling and disease. *Cell* 149(6):1192-1205.

Biography

Eric Jiang was born in Zhaoyuan, China on October 22nd, 1984. He grew up in China, Anchorage, Alaska, and Houston, Texas. He attended Duke University where he received his bachelor of science in Biology in 2007. He joined Duke University's Program in Genetics and Genomics in 2008 and joined the lab of Dr. Joseph Nevins and subsequently, the lab of Dr. Jen-Tsan Ashley Chi.

Publications

Jiang X, Nevins JR, Shats I, Chi JT. E2F1-mediated induction of NFYB attenuates apoptosis via joint regulation of a pro-survival transcriptional program. *Manuscript in submission*.

Rempel RE*, Jiang X*, Fullerton P, Tan T, Ye J, Lau JA, Mori S, Nevins JR, Friedman DR. Utilization of the Eμ-myc Mouse to Model Heterogeneity of Therapeutic Response. *Molecular Cancer Therapeutics*. *In Press*.

Cho AH, Jiang X, Mann DM, Kawamoto K, Robinson TJ, Wang N, McCarthy JJ, Woodward M, Ginsburg GS (2012). A more rapid approach to systematically assessing published associations of genetic polymorphisms and disease risk: type 2 diabetes as a test case. *Comparative Effectiveness Research*. 2,1-20.

Brady SM, Zhang L, Megraw M, Martinez NJ, Jiang E, Yi CS, Liu W, Zeng A, Taylor-Teeples M, Kim D, Ahnert S, Ohler U, Ware D, Walhout AJ, Benfey PN (2011). A stele-enriched gene regulatory network in the Arabidopsis root. *Molecular Systems Biology*. 7:459.

* co-first authors